PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/02, 15/03, 15/04, 15/05, 15/06, 15/08, 15/74, 15/79

A1

(11) International Publication Number: WO 00/06715

(43) International Publication Date: 10 February 2000 (10.02.00)

(21) International Application Number: PCT/US99/16297

(22) International Filing Date: 26 July 1999 (26.07.99)

(30) Priority Data:
60/094,294 27 July 1998 (27.07.98) US

(71) Applicant *(for all designated States except US)*: GENOTYPES INC. [US/US]; Suite B, 61 Airport Boulevard, South San Francisco, CA 94080 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: HITZEMAN, Ronald [US/US]; 15 Banff Way, Pacifica, CA 94044 (US). CHISOLM, George, E., IV [US/US]; 864 Overlook Court, San Mateo, CA 94403 (US).

(74) Agents: TRECARTIN, Richard, F. et al.; Flehr Hohbach Test Albritton & Herbert LLP, Suite 3400, 4 Embarcadero Center, San Francisco, CA 94111-4187 (US).

(81) Designated States: AU, CA, JP, US, European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE).

Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(54) Title: AUTOMATIC EUKARYOTIC ARTIFICIAL CHROMOSOME VECTOR

(57) Abstract

The present invention is directed to novel recombinant nucleic acids for introducing yeast chromosomal elements into the genomes of bacteria. The invention provides methods to convert the modified bacterial genomes into artificial yeast chromosomes by fusing the bacteria with yeast that linearize the modified bacterial genomes, to produce artificial chromosomes.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

AUTOMATIC EUKARYOTIC ARTIFICIAL CHROMOSOME VECTOR

FIELD OF THE INVENTION

The present invention relates to novel recombinant nucleic acids and methods for their use to introduce a prokaryotic genome or other valuable DNA into a eukaryotic cell as a circular molecule that is then converted to a linear automatic eukaryotic artificial chromosome within the eukaryotic nucleus and as a circular chromosome within a membrane bound, extranuclear element. Such added prokaryotic genomes or other valuable DNAs should add new functions to the eukaryotes or allow them to be selected for using this hybrid system.

BACKGROUND OF THE INVENTION

Evolutionary theory proposes that mitochondria and plastids originated by engulfment or cell fusion of prokaryotes by eukaryotes. As this relationship evolved, the size of the bacterial DNA genome decreased and the functions of genes lost from the bacterial genome were assumed by the eukaryotic chromosome (Cavalier-Smith. (1987) *Ann. NY Acad. Sci.* 503:55-71). Support for this theory is found in the fungus, *Geosiphon pyriforme*, which contains in its hyphal system cyanobacteria belonging to the genus, *Nostoc*, but which retain the capacity for autonomous growth and replication (Mollenhauer. (1992) *Geosiphon pyriforme*, In *Algae and Symbiosis*. Biopress Ltd., Bistol, pp. 339-351). Additional support is found in algae which have plastids containing DNA that has a significant level of homology and similar gene organization to cyanobacteria but the plastids have lost most of the cyanobacterial genes to the cell nucleus (Douglas. (1994) Chloropast Origins and Evolution, In *Molecular Biology of Cyanobacteria*, vol. 1 (Bryant, ed.) Kluwer Academic Publishers, Boston, pp. 91-118). The reason and mechanism for the relocation of a large proportion (greater than 90%) of the bacterial genes to the nucleus are unknown (Valentin *et al.* (1992) Phylogenetic origin of the plastids, In *Origins of Plastids* (Lewin, ed.) Chapman and Hall, New York, pp.193-221).

It would therefore be useful to develop a system to introduce an entire prokaryotic genome into a eukaryotic organism and to study the interactions of the two genomes and the effect this has on both organisms. Preferably, such a system would permit both nuclear and extra-nuclear localization of the bacterial genome. This system also would provide a model for the evolution of mitochondria, chloroplasts, and other plastids.

The present invention presents new technology that can be used to transfer entire bacterial chromosomes into yeast or other eukaryotic organisms in such a manner that they become functional linear artificial chromosomes and, furthermore, may become compartmentalized bacteria or organelle-like structures. The bacterial chromosome will be expressed partially in the nucleus and in the bacterial organelle and provides new and useful pathways to the eukaryote host immediately after formation of the hybrid cell or after selections for specific desired functions normally done by the prokaryote alone as well as new functions. These new vectors and methods additionally provide a means to efficiently introduce very large segments of DNA into eukaryotic cells without extracellular manipulation.

SUMMARY OF THE INVENTION

The present invention provides compositions and methods for transferring an entire prokaryotic genome or other DNAs into a eukaryotic organism. In one aspect, the invention provides a recombinant expression system to introduce an endonuclease gene with a rare cleavage site into a eukaryotic organism as well as using endonuclease(s) already present in the eukaryote. In another aspect, the invention provides circular recombinant nucleic acids that are converted by the endonuclease to automatic, eukaryotic artificial chromosomes. The invention also provides a recombinant nucleic acid for converting a prokaryotic genome into a eukaryotic chromosome. In yet another aspect, the invention provides methods for introducing converted bacterial genomes and circular DNAs into a eukaryotic organism. Finally, the invention also provides methods for selecting eukaryotic organisms comprising the modified bacterial chromosome or other DNAs, as well as the selectable addition of new valuable functions from the prokaryotes or other DNAs being added to the eukaryotic cell.

BRIEF DESCRIPTION OF THE DRAWINGS

Figure 1 shows plasmid GET735 and the origin of the components used for its construction: Yeast *LEU2* 5' region, pUC19, *LEU2* 3' region, *URA3* promoter and terminator, and the *HO* restriction enzyme gene operatively linked to the pGAL10 promoter and alcohol dehydrogenase (ADHt) 3' transcription terminator.

Figure 2 shows the nucleotide sequence of two synthetic dual *HO* sites (Panel A) (SEQ ID NOs: 1-2) and PI-*SceI* (Panel B) (SEQ ID NOs: 3-4) endonuclease sites. Selected restriction enzyme sites are shown. The sites of endonuclease cleavage are marked with arrows (4 bp 3' extensions for both the *HO* and the PI-*SceI* sites). GN numbers refer to the synthetic DNAs (size given in bases) used to produce the double stranded DNA. For *HO* sites, although only 24 bases are required for endonuclease recognition and cleavage, additional DNA up to the *Bgl*II sites is also homologous to DNA normally adjacent to the *HO* site *in vivo*.

Figure 3 shows the pYAC (GET149) and pAYAC (GET774) plasmids to scale with selected restriction sites given in Panel A and Panel B, respectively. The actual size of each plasmid is given in base pairs. The white boxed-in sequences show functional units. Single lines are pBR322 sequence. Blacked boxed-in regions are yeast sequences and *SUP4°* is an ochre suppressor tRNA gene from yeast.

Figure 4 shows the larger pYAC (GET860)) (Panel A) and pAYAC (GET856) (Panel B) with the added yeast DNA as smaller white blocked-in regions. Yeast DNA is also shown as solid black blocks. Functional units are shown as the wide white blocks. Lines in the circular plasmid represent pBR322 DNA.

Figure 5 shows the result of the two Southern blots using biotin detection (BLUEGENE) of bands hybridizing with pBR322 DNA that has been nick translated/biotin labeled. Various pBR322 DNA digests were used as molecular weight markers in lanes 4, 9, 10, and 17. Dark bands on the nitrocellulose correspond with DNA fragments that hybridize with the labeled probe. Lanes 5, 7, 11, and 14: GYT3678. Lanes 6 and 8: GYT3677. Lanes 12 and 15: GYT3693. Lanes 13 and 16: GYT3695. The extracted yeast DNA was restriction cut using the enzymes shown at the bottom of the gel with arrows showing which enzymes were used on the DNA in each lane. *NotI*/*Bam*HI shows a combination of two restriction enzymes used. Smaller arrows designate band locations.

Figure 6 shows pYAC + pyrD Ori plasmid. The *E. coli pyrD* sequences are shown as boxed regions with an arrow in them showing the direction of transcription in this gene. Other functional units are boxed. The origin of replication which works in *E. coli* is designated. pAYAC plasmids are all constructs other than #1. The "Site" referred to in Constructs #2-#4 is any desired endonuclease cleavage site (e.g. *MATa* HO, *MATα* HO, PI-*SceI* or any other enzyme site) while the selectable marker is any yeast selectable marker (HIS3 is used as an example only).

Figure 7 shows the general structure of elements for genomes to be used as pAEAC's. The general structure of a pAEAC for use in transfer of prokaryotic genomes into eukaryotic cells is shown in Panel A. Panel B shows the general structure of a pAEAC to be used for gene therapy or

the addition of specific gene systems into eukaryotic cells without the incorporation of the prokaryotic genome sequences. The functional elements are identified but are not drawn to scale.

DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

I. Definitions

The term "automatic yeast artificial chromosome" (AYAC) when used herein encompasses a circular recombinant nucleic acid molecule that is converted to a linear yeast chromosome *in vivo* by an endogenously expressed restriction endonuclease. For use in other eukaryotic species this is called an "automatic eukaryotic artificial chromosome" or AEAC. The recombinant nucleic acid carries appropriately oriented sequences that function as telomeres in the eukaryote and sequences that function as centromeres in the eukaryote used, and a replication origin(s) for autonomous replication within yeast or other eukaryotes. The recombinant nucleic acid should contain selectable markers that work both in the prokaryotes and the eukaryotes.

A "recombinant" or "isolated" nucleic acid molecule comprising the various nucleic acid sequences disclosed herein, means a nucleic acid molecule that has been assembled by molecular biological techniques to contain sequences of defined function that are operably linked.

The term "control sequences" refers to DNA sequences necessary for the expression of an operably linked coding sequence in a particular host organism. The control sequences that are suitable for prokaryotes, for example, include a promoter, optionally an operator sequence, and a ribosome binding site. Eukaryotic cells are known to utilize promoters, polyadenylation signals, and enhancers.

A nucleic acid is "operably linked" when it is placed into a functional relationship with another nucleic acid sequence. For example, DNA for a presequence or secretory leader is operably linked to DNA for a polypeptide if it is expressed as a preprotein that participates in the secretion of the polypeptide; a promoter or enhancer is operably linked to a coding sequence if it affects the transcription of the sequence; a ribosome binding site is operably linked to a coding sequence if it is positioned so as to facilitate translation. Generally, "operably linked" means that the DNA sequences being linked are contiguous, and, in the case of a secretory leader, contiguous and in reading phase. However, enhancers do not have to be contiguous. Linking is accomplished by ligation at convenient restriction sites. If such sites do not exist, the synthetic oligonucleotide adaptors or linkers are used in accordance with conventional practice.

II. Compositions and Methods of the Invention

The present invention provides recombinant nucleic acids and methods for their use for the transfer of an entire prokaryotic genome or other useful DNAs into a eukaryotic organism. This

transfer may be used to study the origins of prokaryotic-like organelles, such as mitochondria and chloroplasts, or to add new functions and pathways to a eukaryotic organism such as, but not limited to, photosynthesis, nitrogen fixation, thermal or salt resistance, motility, mixed antibiotic pathways, non-mutated eukaryotic genes (e.g. gene therapy), etc.

The present invention provides a recombinant nucleic acid that functions as an automatic yeast artificial chromosome (AYAC) or more generally an automatic eukaryotic artificial chromosome (AEAC). pAYACs are circular plasmids composed of a yeast selectable marker, an *ARS* sequence (or foreign sequences which have ARS-like functions) for replication in yeast, a centromere that is functional in yeast, at least two inverted *Tetrahymena* or yeast telomeres (or telomeres from other species that function in yeast), one or more rare restriction endonuclease sites between the two inverted telomeres, a prokaryotic selectable marker, and flanking prokaryotic sequences for integration into a bacterial chromosome by homologous recombination. pAYACs also preferably contain a second prokaryotic selectable marker and a prokaryotic replication origin that can be removed prior to integration into the circular DNA of the prokaryotes. Some contain the chosen endonuclease gene expressed using the appropriate eukaryotic transcription/translation DNA signals.

The pAYAC vectors are linearized by restriction enzyme digestion and transformed into a prokaryotic organism where they integrate into a circular prokaryotic chromosome by homologous recombination or by the use of other integration methods. Successfully transformed and recombined prokaryotes are preferably isolated by a selectable marker encoded by the integrated pAYAC vector and, subsequently, undergo protoplast fusion to a yeast strain that produces a specific endonuclease, located in the nucleus, that cuts the integrated pAYAC between the inverted telomeres at the endonuclease(s) sites. Once linearized in the nucleus of a eukaryotic host, the bacterial genome functions as an artificial yeast chromosome. Selectable markers such as drug resistance genes or genes that complement yeast auxotrophies ensure the presence and maintenance of the artificial chromosome. If the bacterial chromosome is too large for maintenance in the yeast as a single molecule, additional pAYAC vectors may be inserted into the circular bacterial genome at various locations to allow the genome to be broken down into smaller yeast chromosomes. The establishment of a fused bacterium as an organelle following protoplast fusion may also occur. However, the establishment of the bacterium as an organelle preferably requires a second protoplast fusion with an identical bacterium that does not contain an integrated pAYAC using a different selectable marker. In either instance, selection for new bacterial enzymes or pathways can then be used to isolate the yeast or eukaryote with desired newly acquired characteristics.

One advantage of this system over existing technologies is that an entire genome from one organism can be easily introduced into a second organism without *in vitro* manipulation. This is

accomplished by first modifying a prokaryotic (e.g. bacterial) genome *in vivo* and introducing the entire genome into a eukaryotic organism, such as, but not limited to, a yeast, by protoplast fusion of both the bacterium and the eukaryote. New functions are then selected for, such as mitochondrial function in *rho*$^-$ or *rho*^{o} yeast (little or no unique mitochondrial DNA and loss of mitochrondrial function), photosynthesis (growth with CO_2, H_2O, light, salts, and minerals), nitrogen fixation (growth using atmospheric N_2 as the sole nitrogen source), growth at high temperatures following thermophilic bacterial fusions, motility, creation of new antibiotics from the combination of bacterial antibiotic pathways from more than one bacterial strain, complementation of a mammalian genetic defect (i.e. gene therapy), etc.

A second advantage of this method is that large DNA segments are easily manipulated and introduced into yeast or other eukaryotes. The length of YACs made by conventional methods are somewhat limited and average about 1 Mbp (Larin, Z., Monaco, A.P., and Lehrach, H. (1996) Generation of large insert YAC libraries, In *Methods in Molecular Biology*, YAC Protocols, vol. 54 (D. Markie, ed.), Humana Press Inc., Totowa, NJ, pp. 1-11). This size constraint results from a required *in vitro* ligation of the DNA insert to the YAC vector ends prior to yeast transformation. Although the technology has improved over the years, much larger pieces (3.5 Mbp) have only been transferred by protoplast fusions (Allshire *et al.* (1987) *Cell* 50:391-403). In the examples provided below, the entire *E. coli* genome (4.7 Mbp) will be introduced into a yeast as a circle which is then automatically (via the rare endonuclease activity in the nucleus) converted to a linear automatic yeast artificial chromosome.

The fused eukaryotic organism is, for example, a fungi, a yeast, a protozoa, a plant, an animal cell, a human cell, or a eukaryotic microrganism. Eukaryotic cell or cell lines, such as, vertebrate, invertebrate or plant cells also can be used. The eukaryotic organism preferably expresses an endonuclease that cleaves at a rare site located between the inverted telomeres of the AYAC or AEAC, thereby, converting it into a linear artificial chromosome. Alternatively, the specific endonuclease can be expressed by the AEAC or AYAC vector sequences, from a genomic element integrated into the host genome, or from an extrachromosomal element such as a plasmid or virus vector (can also be an integrated virus). In either case, the restriction endonuclease is operably linked to eukaryotic control sequences, such as a eukaryotic promoter and transcription termination elements.

The prokaryotic organism is, for example, a eubacterium, a cyanobacterium, an archaebacterium. It may have a specific phenotype; for example, a nitrogen-fixing bacterium, a thermophilic bacterium or a prokaryote that produces antibiotics. Examples of prokaryotes include but are not limited to *Escherichia coli*, *Zymomonas mobilis*, *Azotobacter*, *Rhizobium*, *Streptomyces*, *Synechococcus* PCC6301, and *Anabaena* PCC7120.

Selecting the correct endonuclease recognition site is important to AEAC function. Preferably, specific endonucleases are utilized that are characterized by having long restriction site recognition sequences that occur infrequently, hence, rare. Examples of these endonuclease recognition sites include, but are not limited to, *HO* (24 bp), I-*Sce*I (18 bp), and PI-*Sce*I (31 bp). Before choosing a particular endonuclease recognition site, the endonuclease should be expressed in the bacterium and in the eukaryote organism or cell to ensure that the genomes of these organisms are not digested by the chosen enzyme in a way that is toxic to the organism. For example, at least one *HO* endonuclease site it thought to be present in *E. coli* genomic DNA since production of the yeast *HO* endonuclease is toxic to *recA*$^-$ *E. coli* cells, *recA*$^+$ cells, however, are not affected, probably due to repair of occasional double strand breaks by the *recA* gene product (Kostriken, R. and Heffron, F. (1984) The product of the *HO* gene is a nuclease: Purification and Characterization of the Enzyme, *Cold Spring Harbor Symp. Quant. Biol.* 49:89-96). However, it should be noted that *HO* endonuclease is also reported to cleave non-MAT yeast DNA *in vitro* even though only the *HO* site in MAT is cleaved *in vivo*. Comparison of possible *HO* endonuclease target sequences *in vitro* and *in vivo* indicates that the *HO* endonuclease has a lower specificity *in vitro* (Nickoloff, J.A., *et al.*, (1990) In Vivo Analysis of the *Saccharomyces cerevisiae* HO Nuclease Recognition Site by Site-Directed Mutagenesis, *Mol. Cell. Biol.* 10:1174-1179). These observations might account for the apparent cleavage of *E. coli* genomic DNA. It is also highly likely that in the yeast nucleus (*in vivo*) the *HO* endonuclease will only cleave the pAYAC plasmid containing the *E. coli* genome at the desired *HO* endonuclease sites next to the telomeres. I-*Sce*I endonuclease sites are not present in a mammalian genome (CHO cells, Chinese hamster ovary); however, engineered sites introduced into the genome are cut by added endonuclease and lead to toxicity and death accompanied by high recombination frequencies (Sargent, R.G., Brenneman, M.A., and Wilson, J.H. (1997) Repair of site-specific double-strand breaks in a mammalian chromosome by homologous and illegitimate recombination, *Mol. and Cell. Biol.* 17: 267-277). PI-*Sce*I endonuclease sites are most likely not present in the human genome since PI-*Sce*I endonuclease placed in human cells has little or no effect on viability (Brenneman, M., Gimble, F.S., and Wilson, J.H. (1996) Stimulation of intrachromosomal homologous recombination in human cells by electroporation with site-specific endonucleases, *Proc. Natl. Acad. Sci. USA* 93: 3608-3612).

Any chromosome that is circular and is not susceptible to the chosen endonuclease in its genome may be transferred into the yeast or other eukaryotic organisms (e.g. plants, fungi, or animals). Preferably, the yeast or eukaryotic host must also be resistant to the chosen endonuclease. The endonuclease site should not be present in the eukaryotic genome; except for example, *HO* endonuclease recognition site present in the MAT locus found on yeast chromosome

III (Herskowitz, I., Rhine, J., and Strathern, J. (1992) Mating-type determination and mating-type inter conversion in Saccharomyces cerevisiae, In *The Molecular and Cellular Biology of the Yeast Saccharomyces*, vol. 2 (E.W. Jones, J.R. Pringle, and J.R. Broach, eds.) Cold Spring Harbor Laboratory Press, pp. 583-656). Cutting at this site precedes mating type conversion and is not lethal to the cell. Moreover, in the examples that follow, the *HO* endonuclease was successfully expressed in yeast and used to convert pAYAC vectors into linear chromosomes with mating type switching being a control for active *HO* endonuclease function.

In the pAYAC vectors, the specific rare restriction site is placed directly between the inverted telomeres (see Figures 6 & 7 for alternative conformations). The pAYAC vectors also contain all elements required of a ycast chromosome such as a functional centromere and at least one functional yeast origin of replication. Preferably, at least one eukaryotic and one prokaryotic selectable marker is included. The pAYAC vector also contains DNA from a specific gene of a desired bacterium which is disrupted by the insertion of an origin of replication functional in *E. coli* or other convenient vector and, preferably, a second prokaryotic selectable marker. The bacterial replication origin is, for example, the pBR322 origin, and the prokaryotic selectable marker is, for example, the ApR (ampicillin resistance). These sequences are present only for convenience to manipulate the pAYAC vector as a bacterial plasmid and are removed prior to integration into the bacterial genome. In addition, the absence of the ApR gene is preferred for protoplasting the bacterium and for making stable bacterial mitochondria due to the inhibition of peptidoglycan crosslinking or cell wall formation.

Transformation of the bacterium is performed after the bacterial origin and ApR gene (optional) are removed from the plasmid by restriction digestion. The bacterial origin must be removed only if the *E.coli* replication origin from pBR322 works in the target bacterium. The linearized DNA is integrated into the circular genome of the bacterium by recombination between the disrupted bacterial DNA sequence on the linearized pAYAC and the homologous sequence in the bacterial genome. After selecting for various antibiotic resistances or other markers on the DNA fragment, bacteria containing a recombinant genome are isolated. The phenotype of the disrupted gene in the transformed bacterium verifies integration of the plasmid at the correct site.

The transformed bacterium is then spheroplasted using lysozyme or ampicillin and fused with a yeast protoplast, produced using an enzyme such as a glucanase (e.g. glusulase and zymolyase). The yeast protoplast contains multiple mutations (or, alternatively, no mutations if antibiotic resistance is selected for in yeast using a ycast driven promoter) and expresses an active endonuclease under the control of a yeast constitutive or inducible promoter from a plasmid or integrated DNA. After protoplast fusion the entry of the circular bacterial DNA into the yeast nucleus is selected for using various yeast markers (i.e., *TRP1* and *URA3*) present in the pAYAC

integrated into the bacterial genome. The endonuclease cleaves the recombinant bacterial genome, resulting in the formation of a functional automatic yeast artificial chromosome.

The placement of the entire bacterial chromosome into a yeast nucleus as an artificial chromosome or in the cytoplasm as a bacterial organelle will provide all the information required for bacterial functions. For example, because *E. coli* can grow on xylose as a carbon source and yeast cannot (but do transport it), selection for yeast that utilize xylose as a carbon source can be conveniently used as a marker for expression and functioning of the bacterial genome. Fusion of bacteria containing the AYAC vector with a *rho*$^-$ or *rho*^{o} (mitochondrial function and DNA deficient) yeast will be examined for growth on glycerol/ethanol media which requires complex complementation of many pathways or genes in yeast. The *rho*^{o} complementation will require the bacterial replacement of at least 19-23 genes as well as 27 RNA transcriptional units of the yeast mitochondrial DNA (varies in size from 74-85 kbp) and the possible function of an estimated 215 genes located in the nucleus. (Grivell, L.A. (1995) Nucleo-mitochondrial interactions in mitochondrial gene expression, *Critical Rev. in Biochem. and Mol. Biol.* 30: 121-16). These pathways as well as other individual mutations may be complemented by similar functions supplied by the correct functioning of the bacterial DNA as a yeast chromosome or as a circular chromosome in a membrane bound organelle.

The establishment of a bacterial organelle preferably requires the presence of ampicillin or mutations in the bacterium preventing peptidoglycan cross-linking or synthesis in order to prevent the formation of a cell wall. The transformation and complementation of *rho*$^-$ or *rho*^{o} yeast with *in vitro* isolated wild-type or mutant mitochondria (Pon L., and Schatz, G. (1991) Biogenesis of mitochondria, in *The Molecular and Cellular Biology of the Yeast Saccharomyces*, vol.1 (E.W. Jones, J.R. Pringle, and J.R. Broach, eds.) Cold Spring Harbor Laboratory Press, pp. 333-406) indicates that the bacterial cell may also form a functional organelle. In addition to protoplast fusion to make the desired organism, whole mitochondrial transformation by use of metal projectiles that are shot into yeast cells by a particle gun indicates that such bacteria may be injected into plants, tissues, and cells using this method (T.D. Fox *et al.* (1988) *Proc. Natl. Acad. Sci. USA* 85:7288-7292).

Once the eukaryote-prokaryote hybrids are made, new functions of the eukaryote are noted, characterized, or selected for. Additional evidence of bacterial gene expression is determined using specific inhibitors of prokaryotes that do not affect eukaryotes, such as rifampicin (inhibits RNA synthesis in some bacteria), chloramphenicol (inhibits peptidyl transferase in translation), erythromycin (inhibits translocation in translation), fusidic acid (inhibits elongation in translation), streptomycin (inhibits chain initiation in translation), and tetracycline (inhibits binding of aminoacyl-tRNAs to ribosome).

Technical limitations associated with these hybrid bacterial/eukaryotic systems can be anticipated. For example, restriction modification systems in various bacteria may degrade the yeast or eukaryotic DNA. Therefore, *E. coli* strains that are deficient in restriction and modification systems are preferable. This is evidenced by protoplast fusions of bacterial and human cells (Rassoulzadegan, M., Binetruy, B.and Cusin, F. (1982) High frequency of gene transfer after fusion between bacteria and eukaryotic cells, *Nature* 295:257). Although yeast DNA is not methylated, methylation by several heterologous bacterial enzymes (*dam*, the *Sau*3AI, and the *Sss*I methyltransferases) does not appear to effect viability or inhibit cellular functions (Kladde, M.P., and Simpson, R.T. (1996) Chromosome structure mapping *in vivo* using methyltransferases, *Methods in Enzymol.* 274:214-233). This indicates that pre-treatment of the yeast (and other eukaryotes) with the appropriate methyltransferases preferably makes their chromosomal DNAs resistant to specific restriction enzymes from a bacterium that is restriction competent. This procedure has been successfully employed for cloning of DNA from a bacterium into a cyanobacterium (Elhai, J., and Wolk, C.P. (1988) Conjugal transfer of DNA to cyanobacteria, *Methods in Enzymol.* 167:747-754). Proteases also may be problematic when the bacterial cell lyses during fusion and releases its contents; however, this was not a problem when fusing *E. coli* to human cells, described above (Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Introduction of recombinant vectors into mammalian cells. In *Molecular Cloning, A Laboratory Manual*, 2nd edition vol. 3 (C. Nolan, ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 16.30-16.81).

Differences in expression mechanisms between yeast and bacteria will most likely affect gene expression in the hybrid organism. However, it is expected that both *E. coli* and yeast will recognize the heterologous DNA unless the artificial yeast chromosome is localized, in addition to the nuclear location, within an organelle or membrane bound structure outside of the nucleus. Differences between eukaryotic and prokaryotic gene expression include the fact that transcription and translation are coupled in bacteria but occur in separate compartments in yeast. In addition, eukaryotic nuclei contain complexes of DNA, RNA, and proteins (mainly histones) in structures called chromatin, while bacteria or prokaryotes contain no such structures.

Although many metabolic pathways are shared in common by yeast and bacteria, hybrid protein-protein interactions may, in some instances, be detrimental to the function of these pathways in the hybrid organism. If there is a conflict of the two expression systems within the nucleus, an additional yeast nucleus can be introduced into the hybrid organism or cell by fusion with another yeast spheroplasts which happens during protoplast fusion. Such karyogamic yeast are stable (do not form single nucleii) mitotically unless they are exposed to a pheromone from the opposite mating type and mate to fuse nucleii. To enhance this effect, karyogamy mutants of yeast which show no

nuclear fusion can be used (Marsh, L., and Rose, M.D. (1997) The pathway of cell and nuclear fusion during mating in S. cerevisiae, In *The Molecular and Cellular Biology of the Yeast Saccharomyces*, vol. 3 (E.W. Jones, J.R. Pringle, and J.R. Broach, eds.) Cold Spring Harbor Laboratory Press, pp. 827-888). There are certainly numerous concerns about deleterious interactions as well as lethality effects; therefore such yeast and bacterial hybrid organisms preferably require various selections and mutations as known in the art to obtain the desired function in the chimeric organism.

In addition to bacteria such as *E. coli*, hybrid organisms made with other prokaryotes, such as cyanobacteria (photosynthetic and nitrogen fixation pathways for yeast and other eukaryotes) ((1994) *The Molecular Biology of Cyanobacteria* vol. 1, (ed. D.A. Bryant) Kluwer Academic Publishers, Dordrecht, Boston and London), thermophilic bacteria such as thermophilic archaca (to obtain heat resistant yeast as well as other heat resistant eukaryotes), various combinations of *Streptomyces* (to mix antibiotic pathways in yeast) as well as other simple and complex pathways from other bacteria can be utilized and exploited. *Zymomonas mobilis* (Zhang, M., Eddy, C., Deanda, K., Finkelstein, M., Picataggio, S. (1995) Metabolic engineering of a pentose metabolism pathway in ethanologenic *Zymomonas mobilis*, *Science* 267:240-243) bacterium, which produces high levels of alcohol, may be used to increase alcohol production in yeast. A photosynthetic yeast may be used for thc production of alcohol (for gasohol and other uses) from CO_2, H_20, salts, and minerals using a light and dark reaction and possibly in conjunction with nitrogen fixation from atmospheric N_2. In this example, glycolysis would occur during the dark reaction to produce alcohol from the glucose made by the light reaction (photosynthesis). This will require a complex fermentation engineering system to maximize light and CO_2 uptake during exposure to light and to minimize dissolved O_2 in the absence of light. It will also preferably require inhibition of yeast growth during light exposure, atmospheric gas use, and O_2 release which could be done by putting the α factor gene into the genome of yeast using a light inducible, O_2 inducible, or other inducible promoter to automatically stop *MAT*a yeast from growing during daylight hours (see references concerning yeast mating type). This would bypass the lengthy and costly process currently in use of production of the sugar by corn and the use of this corn sugar to feed yeast which then produce the alcohol. The system, described above, could be used to control the conversion of biomass (glucose) directly to alcohol in a single, light dependent organism. This could greatly reduce the cost of alcohol production for use as gasohol and also reduce pollutants and energy requirements for gasohol production. This type of hybrid yeast is preferably produced using the industrial yeasts that are employed in the existing process.

The production of AEAC-bacterial genomes for plants preferably require different telomeres, centromeres, and selectable markers that function more efficiently in plants. The transfer of bacterial

genomes into plants may be used for the production of nitrogen fixing plants from ones that do not fix nitrogen. This would be especially useful for corn, rice, wheat, and other crops that take their nitrogen out of the soil. Hybrids may be made, for example, by fusing plant seed cells with various cyanobacteria protoplasts containing a pAEAC system. *Azotobacter*, or *Rhizobium* containing a pAEAC system preferably are added to plants using a gun injection technique (see above) or protoplast fusions (R.E. Kingston (1997) Electroporation into Plant Protoplasts, In *Current Protocols in Molecular Biology,* vol. 1 (F. Ausubel, R. Brent, R. Kingston, D. Moore, J. Seidman, J. Smith, and K. Struhl, eds.) John Wiley & Sons, pp. 9.3.2-9.3.3). Preferably plant protoplasts are used that can be selected and grown to produce entire plants (Rhodes, C.A., Pierce, D.A., Mettler, I.J., Mascarenhas, D., and Detmar, J.J. (1988) Genetically transformed maize plants from protoplasts, *Science* 240:204-207). Furthermore, *Rhizobium* chromosomes added to the nuclei of plant legumes preferably are used to eliminate the necessity for reinfection of root cells by the whole bacterium (in itself a valuable invention). In addition, plants can be produced that have improved thermoresistance or salt tolerance by fusion of plant cells with archaebacteria containing pAEACs.

The production of AEAC-bacterial genomes for animals and human cells (AHACs) will require different telomeres, origins of replication, centromeres, and selectable markers that function in these cells to make automatic chromosomes. All of these are of similar DNA size to yeast except for the centromere which for humans is very much larger than yeast centromeres - up to several megabases compared with 125 bp for *Saccharomyces cerevisiae* (Harrington, J. J. *et al.* (1997) Formation of *de novo* centromeres and construction of first generation human artificial microchromosomes, *Nature Genetics* 15: 345-355). The centromeres are composed of repeats of 171 bp α-satellite (alphoid) DNA which have been placed together in arrays of about 1 megabase in BACs (bacterial artificial chromosomes) or cloned as arrays of around 100 kbp from chromosome parts in YACs to make artificial chromosomes in human cells (Grimes, B. and Cooke, H. (1998) Engineering mammalian chromosomes, *Human Mol. Genet.* 7: 1635-1640). These systems suggest the ability to apply our automatic artificial chromosome system to gene therapy. Gene therapy has suffered badly from delivery systems and from the low frequency of cells expressing the needed gene for the necessary length of time (Prince, H. M. (1998) Gene transfer: a review of methods and applications, *Pathology* 30: 335-347). AHACs carrying genomic or cDNAs for the necessary genes to be transferred would be ideal due to their formation of automatic functional chromosomes upon reaching the nuclei of the human cells. AHACs stability in bacteria (see Example 13) would allow for cost effectiveness in vesicular delivery systems (see above ref.). The use of attenuated bacteria which promote their phagocytosis into human cells, could improve AHAC delivery to enough cells in the body. Examples of bacteria that promote their uptake into human cells are the pathogens: *Yersinia pseudotuberculosis*, *Shigella flexneri*, *Listeria monocytogenes*, *Salmonella*

typhymurium, and *Brucella abortus* (Pizarro-Cerda, J. *et al.* (1997) When intracellular pathogens invade the frontiers of cell biology and immunology, *Histol. Histopathol.* 12: 1027-1038). As an alternative to using attenuated bacteria, the proteins that induce engulfment are used to make artificial membrane vesicles containing the AHAC DNA (eg. see mechanisms for *Listeria* : Cossart, P. (1998) Interactions of the bacterial pathogen *Listeria monocytogenes* with mammalian cells: bacterial factors, cellular ligands, and signaling, Folia Microbiol. 43: 291-303).

The following examples are offered for illustrative purposes only, and are not intended to limit the scope of the present invention in any way.

All patent and literature references cited in the present specification are hereby expressly incorporated by reference in their entirety.

EXAMPLES

Methods for bacterial growth, plasmid constructions, and transformation, except where indicated, are described in Sambrook *et al.* (1989) *Molecular Cloning, A Laboratory Manual*, 2nd edition, (C. Nolan, ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.

Yeast growth, transformation, molecular genetics, DNA purification, vector components, *E. coli* molecular biology, DNA sequencing, and polymerase chain reaction (PCR) are described in Ausubel *et al.*, (1997) Current protocols in molecular biology. Greene Publishing Associates and John Wiley & Sons, Inc., New York, NY.

E. coli strains used for DNA component isolation and plasmid constructions are: i) DH5α (*deoR*, *endA1*, *gyrA96*, *hsdR17*($r_K^-m_K^+$) *recA1*, *relA1*, *supE44*, *thi-1*, Δ(*lacZYA-argFV169*), φ80δ*lacZΔM15*, F^-, λ^-) (Hanahan. (1983) *J. Mol. Biol.* 166:557-580) and ii) 294̄ (*F* , *endA*, *hsdR17*($r_K^-m_K^+$), *supE44*, *thi-1*, *relA1(?)*, *rfbD1(?)*, *spoT1(?)*, Δ(*lacZYA-argFV169*), φ80δ*lacZΔM15*, F^-, λ^-).

E. coli K12 strains used for AYAC integration into bacterial chromosomes are: ER1398 {F^-, *endA1*, *hsdR2*($r_K^-m_K^+$), *supE44*, *thi-1*, *relA(?)*, *rfbD1(?)*, *spoT1(?)*, *mcrB1*$^-$} (Kelleher and Raleigh (1991) *J. Bact.* 173:5220-5223) and NM522 {λ^-, *F'*, lacI^{q}Δ(*lacZ)M15*, *proA*$^+$*B*$^+$*/supE*, Δ(*lac-proAB*), thiΔ(*hsdMS-mcrB*)*5*, ($r_K^-m_K^+$ McrBC$^-$)} (Woodstock *et al.*, (1989) *Nucleic Acids Res.* 17:3469-3478).

Saccharomyces cerevisiae, GY5325 (MATα *ura3-52 trp1-Δ63 his3-Δ200 GAL*) was obtained as a spore from the genetic cross of YPH499 (GY5097) (MATa *ura3-52 lys2-801$_a$ ade2-101$_o$ trp1-Δ63 leu2Δ1 his3-Δ200 GAL*) (Sikorski and Hieter. (1989) *Genetics* 112:19-27) with X2180-1B (MATα *SUC2 mal mel gal2 CUP1*) (Mortimer and Contopoulou. (1995) In *Yeast*

Genetic Stock Center Catalog, 8th Edition, Dept of MCB/Division of Genetics, Univ. of Cal., Berkeley, CA). Yeast GY5345 (MATa *ura3-52 trp1-63 ade2-101$_o$ his3-Δ200 leu2-Δ1 GAL*) was obtained after a second cross of the progeny from the first cross above described which yielded GY5325. Yeast GY5328 (MATα *ura3-52 trp1-Δ63 his3-Δ200 leu2*::pGAL10-HO/*URA3 GAL*) was generated from strain GY5325 (described above) by transformation with linear DNA which resulted in the integration of the yeast HO endonuclease gene, under the control of a galactose-inducible promoter, into the chromosomal LEU2 locus. Strain S1799D (α*trp5 his4 ade6 gal2*) (Moss (1964) *Biophys. Res. Comm.* 18:850) was used for genomic DNA isolation.

DNA for Southern blots was purified from yeast using a protoplasting procedure (Ausubel *et al.*, (1987) *Current Protocols in Molecular Biology*, vol. 2. Greene Publishing Associates and John Wiley & Sons, Inc., New York, NY.) in combination with a Qiagen extract and DNA purification procedure (Qiagen Plasmid Mini Handbook, Qiagen, March 1996). DNAs were digested, electrophoresed, and transferred to supported nitrocellulose (BA-S, Optitran) membranes (Schleicher and Schuell). Membranes were hybridized with biotin-labeled pBR322 DNA and visualized using the BLUEGENE nonradioactive detection system (GibcoBRL, cat. no. 18279-018, Gaithersberg, MD).

pYAC5 was from Sigma. All commercially available products and reagents were used according to the manufacturer's protocols, except where indicated.

EXAMPLE 1

Construction of an *HO* Endonuclease Expression System for Yeast

The following details the production of a recombinant DNA construct, GET735 (Figure 1), that is used for the insertion by homologous recombination of the *HO* endonuclease gene of *S. cerevesiae* into the *LEU2* locus of a haploid yeast strain (e.g. strain GY5325).

Plasmid GET735 was constructed by excising the *LEU2* gene from YEp13 (Broach *et al.* (1979) *Gene* 8:121-123) by *SalI/XhoI* digestion (2.2 kbp fragment) and subcloning this fragment into the *SalI* site of pUC19. Two-thirds of the central portion of the *LEU2* structural gene was removed by *Bst*XI and *Bst*EII digestion and replaced by a polylinker: (*Bst*EII)-*NotI-BglII-XhoI-NheI-(Bst*XI) (Note: all restriction enzymes sites in parentheses are destroyed during the procedure). The *URA3* (1.2 kbp) chromosomal fragment from yeast (Fasiolo *et al.* (1981) *J. Biol. Chem.* 56:2324) was modified by addition of (*HindIII*)/*BglII*/*SalI* linkers and cloned into the *XhoI* site of the polylinker. The *HO* gene obtained from plasmid YCp50-*HO* (Herskowitz *et al.* (1991) *Methods in Enz.* 194:132-146) was modified by PCR at its 5' end to contain an *XbaI* site just 5' of the ATG initiation codon. The dual yeast promoter, GAL1-10, fragment was isolated from highly expressed yeast genes by PCR of genomic DNA of yeast strain, S1799D (α *trp5*, *his4*, *ade6*,

gal2). The PCR primers were designed to replace the ATG codons of *GAL1* and *GAL10* with *Eco*RI and *Bam*HI sites. For *HO* gene expression the PCR modified *HO* gene was operably linked to the GAL10 side of the dual promoter and a 326 bp *Hin*dIII fragment containing the alcohol dehydrogenase I (ADHt) transcription termination region. The *HO* endonuclease expression cassette containing the *HO* endonuclease coding region operably linked to the GAL1-10 promoter and alcohol dehydrogenase I transcription termination region (ADHt) was cloned into the *Not*I site of the polylinker adjacent to *URA3*, generating plasmid GET735.

EXAMPLE 2

Production of a yeast strain that expresses *HO* endonuclease

The purpose of this experiment is to produce a yeast strain that expresses the *HO* endonuclease under the control of the inducible GAL10 promoter.

Plasmid GET735 (Figure 1, Example 1) was *HpaI/Sal*I digested to release the fragment containing the *HO* endonuclease expression cassette flanked by *LEU2* 5' and 3' terminal sequences and *URA3*. Yeast strain, GY5325, was transformed with the *HpaI/Sal*I fragment according to standard protocols. Following introduction into GY5325 and transport to the nucleus, homologous recombination mediated by the *LEU2* terminal sequences results in the insertion of the *HO* expression cassette into the *LEU2* locus of the GY5325 genome, producing strain GY5328. *HO* endonuclease expression was demonstrated by galactose-inducible, mating-type changes in GY5328.

EXAMPLE 3

Construction of an *HO* and PI-*Sce*I Restriction Recognition Sites

The purpose of this experiment is to construct plasmids containing *HO* and PI-*Sce*I restriction recognition sites that will be used in the construction of automatic yeast artificial chromosome plasmids.

Figure 2A shows the *MAT*α and *MAT*a *HO* endonuclease recognition sites (Herskowitz *et al.* (1992) *The Molecular and Cellular Biology of the Yeast Saccharomyces*, vol 2 (Jones *et al.*, eds) Cold Spring Harbor Laboraotry Press, pp. 583-656) flanked by *Eco*RI and *Hin*dIII sticky ends. The non-blunt ended, double stranded DNA (Figure 2A; SEQ ID NO:1 and SEQ ID NO:2) was formed by hybridization and ligation of 4 overlapping synthetic DNAs (GN337 (SEQ ID NO:3), GN338 (SEQ ID NO:4), GN341 (SEQ ID NO:5), and GN342 (SEQ ID NO:6)) and cloned into the *Eco*RI and *Hin*dIII sites of pUC118. The twenty-four base pairs required for restriction digestion by the *HO* gene product in each site are underlined. The arrows within the recognition sequence indicate the 4-base 3' overhang produced by *HO* digestion.

Figure 2B shows two PI-*SceI* restriction sites (Gimble, F.S., and Thorner, J., (1993) J. Biol. Chem. 263:21844-21853) flanked by *Eco*RI sticky ends. The non-blunt ended, double stranded DNA (Figure 2B; SEQ ID NO:7 and SEQ ID NO:8) was formed by annealing synthetic oligonucleotides (GN520 (SEQ ID NO:9) and GN521 (SEQ ID NO:10) and cloned into the *Eco*RI site of pUC118. The arrows indicate the 4-base 3' overhang produced by PI-*SceI* endonuclease digestion.

EXAMPLE 4

Construction of a pAYAC Plasmid System

The purpose of this experiment is to construct circular recombinant DNA that replicates as a linear automatic yeast artificial chromosome.

pAYAC (automatic yeast artificial chromosome plasmid; GET774) shown in Figure 3B contains all of the components of pYAC (Figure 3A) but also contains *MAT*a and *MAT*α *HO* endonuclease sites between the *HIS3 Bam*HI sites adjacent to each telomere (*TEL*). pAYAC was made by inserting a *Bam*HI spacer fragment which contained no *Bgl*II sites into the *Bam*HI site between the *HO* restriction sites shown in Figure 2A. The fragment shown in Figure 2A and containing the linker was excised by *Bgl*II digestion and exchanged for the *Bam*HI fragment of pYAC5, thereby replacing the *HIS3* gene fragment. The spacer fragment was removed and replaced by the *HIS3 Bam*HI fragment from pYAC5 to obtain pAYAC (GET774).

EXAMPLE 5

Construction of a Larger pAYAC Plasmid System

Additional DNA was inserted into the plasmids shown in Figures 3A-B, for the production of larger plasmids and, therefore, more stable linear artificial chromosomes (Murray, A.W., and Szostack, J.W. (1983) Construction of artificial chromosomes in yeast. *Nature* 305:189-193) after endonuclease cutting in a yeast nucleus between the two inverted telomeres.

A *Bam*HI chromosomal fragment of 8980 bp was isolated from a partial *Sau*3A library of S1799D (α *trp5, his4, ade6 gal2*) size selected genomic fragments placed in the *Bam*HI site of plasmid YRp7 (Stinchcomb *et al.*, (1979) *Nature* 282:39-43). Using Ade^+/Trp^+ selection (media lacking tryptophan and adenine) and strain YPH499 (MATa *ura3-52 lys2-801$_a$ ade2-101$_o$ trp1*-Δ63 *his3*-Δ*200 leu2*-Δ*1 GAL*) transformed with this pool of DNA from this genomic bank made in *E. coli*, a plasmid containing the 8980 bp fragment which contained the *ADE2* gene was isolated by complementation. A *Bam*HI 8980 bp fragment containing the yeast *ADE2* gene was excised by *Bam*HI digestion and its ends were modified by the addition of linkers to destroy the *Bam*HI/*Bgl*II

sites and introduce *Not*I sites. This DNA was ligated into the *Not*I sites of GET149 and GET774, producing GET860 and GET856, respectively in Figure 4.

EXAMPLE 6

Conversion of circular plasmids into linear yeast artificial chromosomes

To test the for conversion of a circular plasmid into a stable linear chromosome, yeast strain GY5328 (MATα *ura3-52 trp1*-Δ63 *his3*-Δ*200 leu2*::pGAL10-HO/*URA*3 *GAL*) was transformed with plasmids GET149 (pYAC5, Figure 3A), GET774 (pYAC5 with *HO* recognition sites, Figures 3B), GET860 and GET856 (equivalent to GET149 and GET774, respectively, but containing an additional 8980 bp *Not*I fragment encoding the *ADE2* gene and its adjacent chromosomal DNA, Figure 4A-B).

The transformed strains were selected by complementation of the tryptophan auxotrophy present in the parental strain using media containing glucose to prevent expression of the *HO* gene. This insures that all transformed strains contain only circular DNA at the beginning of the experiment. The conversion from circular to linear forms was initiated by incubating the transformed strains in media (0.67% Difco Yeast Nitrogen Base without amino acids (YNB) + 0.5% casamino acids (CAA)) containing 2% galactose as carbon source, thereby inducing the GAL1-10 promoter and *HO* endonuclease expression. The transformed strains incubated in media containing 2% glucose as a carbon source served as negative controls. The transformants were incubated overnight at 30°C in 5 ml roller cultures inoculated at an initial density of 0.1 $OD600_{nm}$/ml. The following morning, the density of the overnight cultures was determined, the cultures were diluted to 1.0 $OD600_{nm}$/ml and serially diluted to 10^{-4} at 10^{-2} increments. An aliquot (0.2 ml) of the 10^{-4} or final dilution of each culture, whether from glucose- or galactose-containing media, was spread onto YNB plates containing 0.5% casamino acids, histidine (36 µg/ml), 2.0% glucose, and 3% agar and incubated for 3 days at 30°C to allow the development of isolated colonies. The phenotypes of the resulting colonies were tested by replica plating on agar plates containing only 0.67% YNB + 2.0% glucose and all but one of the following additions: histidine (36 µg/ml), tryptophan (50 µg/ml), and leucine (12 µg/ml). Mating type was assessed by crossing the single colonies to strains GY5302 (*MAT*a, *lys*2-801$_a$ *GAL*) and GY5303 (*MAT*α *lys*2-801$_a$ *GAL*) and scoring for mating by the formation or prototrophic diploids by replication to agar plates containing only YNB and 2% glucose after an overnight incubation at 30°C to allow mating to occur.

The results shown in Table 1 indicate that, as expected, induction of *HO* endonuclease by galactose growth results in mating type switching from *MAT*α to *MAT*a. In addition, if the transformed plasmid contains an *HO* endonuclease recognition site, it should be converted to a

linear yeast artificial chromosome and the generation of strains that are phenotypically His^- and Trp^+. As shown in Table 1, such colonies were generated from transformants containing either GET774 or GET856 when utilizing galactose as carbon source. Moreover, the frequency of His^- Trp^+ colonies increased 5.2 fold from about 0.63% to 3.3% when using the larger GET856 plasmid. No His^- Trp^+ colonies were isolated from transformants containing GET149 or GET860 under any conditions.

Except for a population of non-mating cells, most of the cells have the parental *MAT*α phenotype when pre-grown in medium containing glucose. However, if the same strains are pre-grown in medium containing galactose, almost complete switching of the mating phenotype is observed (mostly to *MAT*a and more non-maters). The non-maters in the glucose grown cells might result from a low level of *HO* activity produced by low levels of constitutive *HO* expression in these strains. The increased number of non-maters in the galactose grown cells is expected because of the high level of mating switching and possible subsequent matings.

The phenotypes and mating types of all of the His^- Trp^+ colonies were retested. All but one of these 16 colonies retested as His^- Trp^+ Ura^+ Leu^-. For the 3 colonies generated from the transformants containing GET774 DNA, two were non-maters and one was *MAT*a (all possible switched phenotypes). For the colonies derived from the transformant containing GET856 DNA there were 3 non-maters, 7 *MAT*a and 2 *MAT*α (10 out of 12 are possible switched phenotypes).

EXAMPLE 7

Genetic evidence for Linearity of AYAC DNA

These experiments were performed to demonstrate that plasmids containing *HO* endonuclease recognition sites are converted to linear AYACs following induction of the *HO* gene. Genetic stability and bacterial transformation efficiency were used as surrogate markers for linearity because it has been shown that circular DNA is more stable and transforms bacteria more efficiently than linear DNA.

The stability of 5 of the putative linear AYACs (His^- Trp^+) were assessed relative to one circular pYAC (His^+ Trp^+) derived from the GET860. The colonies to be tested were grown for two days at 30°C in rich complete YEPD medium (1% Yeast Extract, 2% Bacto Peptone and 2% glucose) in roller tube cultures. The density of the resulting cultures was determined spectrophotometrically at 600 nm. The cultures were diluted to 1.0 OD_{600nm}/ml and serially diluted at 10^{-2} increments to 10^{-4}. A 0.2 ml aliquot of the 10^{-4} final dilution of each culture was spread onto duplicate YEPD-3% agar plates and incubated for 3 days at 30°C to allow isolated colonies to develop. The phenotypes of the resulting colonies were tested by replica plating on 0.67%

Table 1

AYAC Plasmid Test

Plasmid	*HO* Sites	Carbon Source	Total	Colony Phenotype H+,W+	H−,W−	H−,W+	L−	Mating Type *MAT*a	*MAT*α	None	%H−,W+	Isolated DNA in strain GY5345
GET149	-	Glu	109	109	0	0	109	0	89	20	0.00%	
"	-	Glu	124	124	0	0	124	0	99	25	0.00%	
GET774	+	Glu	106	106	0	0	106	0	93	13	0.00%	
"	+	Glu	88	88	0	0	88	0	75	13	0.00%	
GET860	-	Glu	173	173	0	0	173	0	169	4	0.00%	
"	-	Glu	128	128	0	0	128	0	119	9	0.00%	
GET856	+	Glu	148	148	0	0	148	0	138	10	0.00%	
"	+	Glu	168	168	0	0	168	0	156	12	0.00%	
GET149	-	Gal	120	116	4	0	120	75	0	45	0.00%	
"	-	Gal	85	85	0	0	85	63	2	20	0.00%	
GET774	+	Gal	238	234	1	2	238	177	5	56	0.84%	
"	+	Gal	234	233	0	1	234	164	0	70	0.43%	

Table 1

(continued)

AYAC Plasmid Test

Plasmid	*HO* Sites	Carbon Source	Total	Colony Phenotype H+,W+	H-,W-	H-,W+	L-	Mating Type *MAT*a	*MAT*α	None	%H-, W+	Isolated DNA in strain GY5345
GET860	-	Gal	270	270	0	0	270	168	43	59	0.00%	**GYT3678**
"	-	Gal	243	242	1	0	243	181	35	27	0.00%	
GET856	+	Gal	191	185	0	6	191	109	28	54	3.14%	**GYT3677**
"	+	Gal	202	195	0	7	202	116	38	48	3.47%	

Notes: Yeast Strain GY5328

H=histidine, W=tryptophan, and L=lucine

None=non-mater

Glu=glucose, and Gal=galactose

WO 00/06715

PCT/US99/16297

YNB/0.5% casamino acids/2.0% glucose agar plates with and without tryptophan (50 µg/ml). Strains cured of extrachromosomal DNA (either circular plasmid or linear AYAC) will have a Trp^- phenotype. The circular pYAC plasmid (GET860) cured at an average rate of 3.7% while the five His^- Trp^+ colonies derived from GET856, cured at average rates of 59.9%, 55.5%, 75.1%, 54.7%, and 55.9% respectively, consistent with an His^- Trp^+ phenotype containing a linear chromosomal form.

Transformation efficiencies of *E. coli* strain DH5α were determined using DNA prepared from strains GYT3653 (derived from GY5328 transformed with GET 860, circular pYAC, Figure 4A) and strain GYT3650 (derived from GYT5328 transformed with GET856, a putative linear AYAC, Figure 4B). Transformation of yeast strain GY5345 (*MATa trp1*-Δ63 *ura3-52 ade2-101*, *his3*-Δ*200 leu2*-Δ*1*) served as a positive control. Linear DNA should not transform *E. coli* to ampicillin resistance but both linear and circular DNAs should transform yeast to Trp^+.

As expected, DNA from GYT3653 efficiently transformed *E.coli* while the DNA from GYT3650 generated only a few colonies (less than 5% of the number seen with GYT3653 DNA). Analysis of the plasmids contained in the *E. coli* transformants using restriction endonuclease digestion and agarose gel electrophoresis confirmed that all transformants from the GYT3653 DNA contained intact circular GET860 DNA while all of the transformants from GYT3650 DNA contained a plasmid that appeared to have no relationship to the pAYAC plasmids used in this study. This analysis was repeated and yielded the same results.

DNAs from GYT3653 and GYT3650 were able to transform yeast to Trp^+. A slightly higher number of transformants were produced with GYT3653 DNA but it was not possible to normalize the DNA concentration prior to transformation. Phenotypic analysis of the yeast transformants indicated that all colonies were Ade^+ Ura^+. Colonies generated from GYT3653 DNA were also His^+ while colonies generated from GYT3650 DNA were His^-. These results indicate that the putative linear AYAC DNA from GYT3650 contains all the portions of the basic pYAC DNA (i.e. *TRP1 URA3*) in addition to the added *ADE2* gene, but lacks the *HIS3* gene which is removed by *HO* endonuclease digestion.

The *E. coli* transformation results suggest the possibility that the original colony containing the putative linear AYAC from GET856 (GYT3650) might also contain a low level of a contaminating circular plasmid. To test this we purified DNA from colonies generated by transforming yeast strain GY5345 with DNA from GYT3650 (GYT3677 - putative linear AYAC, Table 1) and GYT3653 (GYT3678 - circular GET860 control, Table 1). This DNA was used to transform *E. coli* strain DH5α to amplicillin resistance. Similar results to those described above were obtained and analysis of the isolated plasmids again confirmed that transformants from GYT3678 (GET860, circular control) contained intact GET860 DNA and transformants from

GYT3677 (putative linear AYAC from GET856) contained a plasmid of unknown origin. DNA from both GYT3677 and GYT3678 was used to transform yeast strain GY5345 and results similar to those described above were obtained. In addition, DNA from *E. coli* transformants generated with GYT3677 and GYT3678 DNA was used to transform yeast strain GY5345. *E. coli* DNA derived from GYT3678 (GET860 circular control) was able to transform yeast to Trp^+ (selection marker) and Ade^+ (Ade^+ and Ade^- colonies are white and red, respectively). *E. coli* DNA derived from GYT3677 (putative linear AYAC from GET856) failed to generate Trp^- colonies, suggesting that this DNA does not contain either a functional *TRP1* gene or a yeast origin of replication.

EXAMPLE 8

Physical evidence for Linearity of AYAC DNA

For Southern analysis, DNA from yeast strains GYT3677 and GYT3678 (Table 1) was purified, *Not*I digested, electrophoresed, transfered to nitrocellulose and hybridized to a biotin-labeled probe. The Southern blot of the *Not*I GYT3678 DNA, Figure 5, lane 5, yields a single band of approximately 11.3 kbp which is consistent with the transformed GET860 being in a circular form. In contrast, analysis of GYT3677 DNA, Figure 5, lane 6, yields bands of >3.6 kbp and >5.9 kbp which is the pattern expected for a linear AYAC with remodeled telomeres resulting in greater DNA fragment length.

*Not*I and *Bam*HI digestion of GYT3678 DNA yielded 3.5 kbp and 5.9 kbp fragments (Figure 5, lane 7) due to removal of the *HIS3* gene from between the two telomeres in GET860 which is expected from a circular molecule. However, *Not*I-*Bam*HI digestion of GYT3677 DNA yielded the expected >3.6 kbp and >5.9 kbp fragments (Figure 5, lane 8). This is consistent with the automatic formation of a linear AYAC in GY5328 transformed with GET856 and grown in the presence of galactose to induce *HO* expression. The results indicate that plasmid GET856 forms a linear AYAC in the presence of an *HO* endonuclease in the yeast nucleus.

EXAMPLE 9

Effect of *HO* recognition site DNA on Telomere Function

This experiment was performed to evaluate the effect of the *HO* site DNA at the ends of the telomeres in construct GET856 on the stability of the linear AYAC DNA and to determine if the low level of His^- Trp^+ transformation observed after *HO* endonuclease induction was due to increased instability of the linear AYAC DNA caused by the *HO* recognition site sequences.

GET860 and GET856 DNA and were linearized by *Bam*HI restriction endonuclease yielding identical linear AYAC molecules except for the presence of the *HO* endonuclease sites on the ends of the telomeres of GET856 (compare Figure 4A and 4B). Both linear DNAs were

transformed into yeast strain GY5345. Selection for Trp$^+$ produced only white colonies indicating co-transformation to Ade$^+$. Several transformants for both linear DNAs were tested for stability and their phenotypes were analyzed. All the transformants were Ura$^+$ His$^-$ as would be expected from transformation with a linear molecule. About twice as many colonies were generated from the GET856 (with *HO* sites) linear DNA than from the GET860 (no *HO* site control) but this most likely represents differences in initial DNA concentration. Transformants were scored for stability by a sectoring assay. Transformants were inoculated into YEPD medium and were grown for 48 hours at 30°C in 5 ml roller tube cultures. The density of the resulting cultures were determined spectrophotometrically at 600 nm. The cultures were diluted to a density of 1.0 OD_{600nm}/ml and serially diluted at 10^{-2} increments to 10^{-4}. A 0.2 ml aliquot of the 10^{-4} dilution of each culture was spread onto YEPD-3% agar plates and incubated for 3 days at 30°C to allow colony development. Stability was assessed by scoring colony phenotype with respect to the *ADE2* locus: i) colonies that had completely lost the linear AYAC were red, ii) colonies that lost the DNA after being plated on YEPD plates generated a sectored red and white colony and iii) colonies that retained the linear AYAC DNA are white.

The results of the sectoring assay is shown in Table 2. The three sections of Table 2 represent the three separate identical assays that were performed to evaluate transformant stability. As shown in Table 2, top section, the circular (Form P Table 2) control plasmids, GET860 and GET856, are very stable in strain GY5345, generating only about 1% loss in this assay. Linear (L) DNA transformants from both GET860 and GET856 are both relatively unstable with about 10-17% loss of plasmid and 13-25% sectoring. The exception is GET856 (A) which is very stable and appears to behave like a circular molecule. Because GET856(A) is His$^-$, this colony could not have resulted from contamination by uncut plasmid DNA in the original transformation.

Testing of additional transformants with linear AYACs yielded the following results (Table 2, middle and bottom sections). First, all transformants from linear GET860 DNA produced unstable Ade phenotypes with similar rates of sectoring (between 10 and 21%). Second, of the 8 transformants analyzed from linear GET856 DNA, 3 produced unstable Ade phenotypes while the other 5 transformants behaved like circular molecules that remained very stable in this assay. Third, the 3 unstable transformants from GET856 behaved just like all the *Bam*HI digested GET860 DNA transformants, suggesting that the GET856 DNA is capable of forming a complete YAC just like GET860 DNA. Fourth, the data derived here further verifies the linear nature of the AYAC found in the His$^-$ Trp$^+$ transformants generated after *HO* endonuclease expression identified in Table 1.

Table 2

Test of AYAC Linear DNAs

	DNA Form	*HO* Sites	Ade Phenotype Total	Red	Sect.	%Red	%Sect.	%R+S	Avc %R+S
GET860(A)	P	-	203	1	4	0.49%	1.97%	2.46%	
GET860(B)	P	-	254	1	1	0.39%	0.39%	0.79%	
GET856(A)	P	+	214	1	1	0.47%	0.47%	0.93%	
GET856(B)	P	+	210	2	0	0.95%	0.00%	0.95%	
GET860(A)	L	-	183	27	24	14.75%	13.11%	27.87%	
GET860(B)	L	-	216	22	54	10.19%	25.00%	35.19%	
GET856(A)	L	+	175	0	2	0.00%	1.14%	1.14%	
GET 856(B)	L	+	154	27	34	17.53%	22.08%	39.61%	
GET860(A)	L	-	420	76	58	18.10%	13.81%	31.90%	29.59%
"	L	-	385	53	52	13.77%	13.51%	27.27%	
GET860(B)	L	-	570	92	105	16.14%	18.42%	34.56%	36.79%
"	L	-	510	87	112	17.06%	21.96%	39.02%	
GET856(C)	L	+	323	2	4	0.62%	1.24%	1.86%	2.77%
(GYT3693)	L	+	380	12	2	3.16%	0.53%	3.68%	
GET856(D)	L	+	423	0	0	0.00%	0.00%	0.00%	0.00%
"	L	+	385	0	0	0.00%	0.00%	0.00%	
GET856(E)	L	+	340	14	61	4.12%	17.94%	22.06%	19.96%
"	L	+	336	12	48	3.57%	14.29%	17.86%	
GET856(F)	L	+	298	0	0	0.00%	0.00%	0.00%	0.00%
"	L	+	249	0	0	0.00%	0.00%	0.00%	
GET856(G)	L	+	420	68	79	16.19%	18.81%	35.00%	33.92%
"	L	+	469	79	75	16.84%	15.99%	32.84%	
GET856(H)	L	+	274	1	0	0.36%	0.00%	0.36%	0.31%
"	L	+	393	1	0	0.25%	0.00%	0.25%	

Table 2

(continued)

Test of AYAC Linear DNAs

	DNA	*HO*	Ade Phenotype						Ave
	Form	Sites	Total	Red	Sect.	%Red	%Sect.	%R+S	%R+S
GET860(C)	L	-	698	69	71	9.89%	10.17%	20.06%	19.05%
"	L	-	937	47	122	5.02%	13.02%	18.04%	
GET860(D)	L	-	423	60	78	14.18%	18.44%	32.62%	30.56%
"	L	-	386	72	38	18.65%	9.84%	28.50%	
GET860(G)	L	-	482	72	68	14.94%	14.11%	29.05%	28.91%
"	L	-	476	73	64	15.34%	13.45%	28.78%	
GET856(G)	L	+	380	63	94	16.58%	24.74%	41.32%	43.26%
(GYT3695)	L	+	365	80	85	21.92%	23.29%	45.21%	

Yeast Strain GY5345

Red=Ade^- colony, Sect.=Sectored Colony (Red Ade^- and White Ade^+)

Form: P=Plasmid, L=Linear

The generation of stable circle-like molecules from the GET856 DNA was further investigated. DNA was purified from one of the stable GET856 transformants and efficiently transformed *E. coli* to ampicillin resistance. The efficiency of the transormation was the first indication that the GET856 DNA had recircularized. Analysis of the isolated plasmid DNA, along with DNA sequencing, indicated that the GET856 DNA had recircularized and undergone a recombination event that eliminated both the *HO* endonuclease cleavage sites, one entire telomere and most of the second telomere.

The results of these studies indicate that the pAYAC is capable of forming a linear molecule in yeast with *HO* site sequences on the ends of the telomere units. However, the presence of these sequences also allows for some degree of recircularization which results in highly stable yeast transformants. However, no stable transformants were identified from the analysis of plasmid stability from the transformants generated in Table 1. This might be due to the differences in the structures of the *HO* recognition site sequences on the ends of the telomere units in each case; for example sites versus cleaved, partial sites. The data also suggest that the desired endonuclease cleavage site(s) are preferably placed adjacent to the telomeres with no intervening sequences to promote the most efficient formation of linear AYACs.

Lanes 10-17 of Figure 5 show a second Southern blot with lanes 10 and 17 being pBR322 standards. GYT3678 (Table 1) as well as GYT3693 and GYT3695 (Table 2) were analyzed in lanes 11-13 and 14-16, respectively. Remember the plasmid DNAs (Figure 4A and 4B) from Table 2 were cut with *Bam*HI before transformation of strain GY5345. Both can now form YACs as has been shown before (Burke *et al.* (1987) Cloning large DNA segments of exogenous DNA into yeast by means of artificial chromosome vectors. *Science* 236:806-812). Again GYT3678 (used on GEL1 and see Table 1) containing GET860 is again shown to be a circle when digested with *Not*I and *Bam*HI giving the expected 3.6 and 5.9 kbp bands (lane 11) and when cut with *Not*I and *Nco*I giving 1.5 and 9.8 kbp bands (lane 14, also see Figure 4). GYT3693 (Table 2) shows the cutting expected for a circular DNA that has lost the *Bam*HI sites. This DNA is cut into 2 fragments when digested with *Not*I and *Bam*HI giving one fragment at 9.5 kbp which is detected in lane 12. GYT3693 (Table 2) DNA was also digested with *Not*I and *Nco*I (see Figure 4B) and is shown in lane 15 to contain 1.5 and 8.0 kbp fragments which hybridize with pBR322 DNA. This is consistent with it being a circular DNA lacking a *Bam*HI site between the telomeric DNA (also based on stability). This aberrant form of GET856 transformed after digestion with *Bam*HI is discussed earlier in the presentation of Table 2. These results suggest that complete, intact *HO* recognition sites on the ends of the telomeres inhibits AYAC formation. However, GYT3695 DNA (Table 2) shows the correct patterns for a properly formed linear AYAC when GET856 is *Bam*HI digested and transformed into GY5345. *Bam*HI and *Not*I digestion produced bands of >3.6 and >5.9 kbp (Figure 5, lane 13) while

*Nco*I and *Not*I digestion produced bands of >5.9, >2.1, and 1.5 kbp (Figure 5, lane 16). Again, the fragments that are slightly larger than the expected sizes are due to remodeling of *Tetrahymena* telomeres after exposure of their ends in the yeast nucleus. These assays were repeated in various other yeast strains and yielded the same results (data not shown).

EXAMPLE 10

Construction of Vectors for Converting a Bacterial Genome into an Automatic Yeast Artificial Chromosome

Figure 6 shows the structure of recombinant DNA plasmids designed to convert a bacterial genome into a large automatic yeast artificial chromosome plasmid (pAYAC). All four constructs contain the same backbone of functional and selectable units for expression and replication in both i) *E. coli* as either a circular plasmid or, after *Kpn*I linearization, as an integrated DNA and ii) yeast, *Saccharomyces cerevisiae*, as either a circular plasmid or after restriction cleavage *in vivo* as a linear chromosome. The plasmids shown in Figure 6 also contain bacterial target sequences, from the *E. coli pyrD* gene in this case, to allow the plasmids to be integrated into the bacterial genome. To function as pAYACs, the plasmids must also contain sites for endonuclease digestion *in vivo* in one of the three configurations depicted (Figure 6, constructs #2-#4) between the telomeres. As mentioned before the "Site" in Figure 6 is for any desired endonuclease (e.g. *MATa HO*, *MATα HO*, PI-*SceI*, etc.) and the selectable marker can be any yeast gene (e.g. HIS3)

First pYAC5 was modified by introduction of the *Cla*I/*Sal*I fragment of pBR322 encoding the tetracycline resisitance gene (TetR), yielding pYAC5+Tet. The *Aat*II/*Nde*I fragment containing the AmpR and a bacterial origin of replication and flanked by *E. coli pyrD* sequences was generated by first using PCR to isolate a portion of the *pyrD* gene of *E. coli* DH5α. The first PCR primer corresponded to nucleotides +1 to +17 (+1 being the A of the ATG initiation codon) of the *pyrD* gene and was designed to introduce *Eco*RI/*Aat*II sites at the 5' terminus of the PCR product. The second primer corresponded to *pyrD* gene bases +446 to +469 and contained *Nde*I/*Hin*dIII sites at its 5' terminus. These two primers were used with DH5α genomic DNA to generate a fragment containing 45% of the *pyrD* structural gene and lacking protein coding sequence from both termini. This PCR product was cloned as an *Eco*RI/*Hin*dIII (455bp) fragment in pUC118. An *Aat*II/*Nde*I fragment from pBR322, containing the AmpR gene and an origin of replication that is functional in *E. coli*, was introduced into a unique *Kpn*I site within the *pyrD* sequences (nucleotide +252). The origin of replication was previously converted to a *Kpn*I fragment while eliminating its *Aat*II and *Nde*I sites. The entire *Aat*II/*Nde*I fragment (2213 bp) fragment containing

the AmpR gene and replication origin flanked by *pyrD* sequences was exchanged into pYAC5+Tet, yielding pYAC+pyrD Ori (Figure 6, Construct #1).

Constructs #2-#4 are identical to Construct #1 but have either one or two rare endonuclease restriction sites between the two inverted *Tetrahymena* telomeres. These constructs will be made by replacing the unique *Xho*I fragment from pYAC+pyrD Ori (Construct #1) that contains the inverted telomeres with *Xho*I fragments containing the desired endonuclease cleavage site structure between the telomeres.

EXAMPLE 11

Conversion of a Bacterial Chromosome into an Automatic Yeast Artificial Chromosome

To convert a bacterial chromosome into an automatic yeast artificial chromosome, Constructs #1-#4 (Figure 6) will be cut with *Kpn*I to remove the AmpR gene and *E. coli* replication origin. The remaining, larger fragment will be used to transform, for example, *E. coli* K12 strain, ER1398 {*F*⁻, *endA1*, *hsdR2*(r_K^- m_K^+), *supE44*, *thi-1*, *relA?*, *rfbD1?*, *spoT1?*, *mcrB1*⁻} and *E. coli* K12 strain NM522 {λ^-, *F'*, *lacIq* Δ(*lacZ*)*M15*, *proA*⁺*B*⁺/*supE*, Δ(*lac-proAB*), *thi*Δ(*hsdMS-mcrB*)*5*, (r_K^- m_K^- McrBC⁻)} to tetracycline resistance by insertion of the entire fragment into the bacteria chromosome. Integration into the bacterial chromosomal *pyrD* gene by homologous recombination requires a single cross-over event. This results in the insertion of the entire fragment and a tandem duplication of two defective *pyrD* genes. The transformants will be Tet resistant, Ura⁻ and Amp sensitive. The stability of the two inverted telomeres will be confirmed by PCR, using primers that hybridize to unique DNA sequences on opposite sides of the telomeres. Alternatively, a primer that hybridizes to the unique sequences of the endonuclease cleavage site used in combination with primers that hybridize to unique sequences flanking the telomere may be used.

An alternative method for integration of the pAYAC plasmid into the bacterial genome would be to use the FLIRT system (Huang, L-C., *et al.*, (1997) Convenient and reversible site-specific targeting of exogenous DNA into a bacterial chromosome by use of the FLP recombinase: the FLIRT system. *J. Bacteriol.* 179:6076-6083). The *E. coli pyrD* target sequences would be replaced by a single 34 bp FLP recombinase recognition site (Broach, J.R. and Volkert, F.C. (1991) Circular DNA plasmids of yeasts, In *The Molecular and Cellular Biology of the Yeast Saccharomyces*, vol. 1 (E.W. Jones, J.R. Pringle and J.R. Broach, eds.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 297-331). The pAYAC plasmid would have the bacterial origin of replication and the AmpR gene removed as above but the rest of the plasmid would be recircularized by ligation prior to transformation into the bacteria. Expression of the yeast FLP recombinase in the bacteria would result in the integration of the pAYAC into the

bacterial genome which also would contain a single 34 bp FRT recombination site. This site could be introduced by the use of a transposon as described for the FLIRT system or could be integrated into the phage lambda attachment site using the lambda integrase gene product.

Yeast strains GY5328 (*HO* endonuclease gene under galactose control, *MATα ura3-52* trp1-Δ63 *his3-Δ200 leu2*::pGAL10-*HO*/*URA3 GAL*) and GY5097 (YPH499 = *MAT*a *ura3-52* *lys2-801$_a$ ade2-101$_o$ trp1*-Δ63 *his3-Δ200 leu2-Δ1 GAL*) have been made *rho*^{o} (essentially no mitochondrial DNA or functional mitochondria) using a standard ethidium bromide procedure (Fox, T.D., Folley, L.S., Mulero, J.J., McMullin, T.W., Thorsness, P.E., Hedin, L.O., and Costanzo, M.C. (1991) Analysis and manipulation of yeast mitochondrial genes, *Methods of Enzymol.* 194: 149-165). These strains will be converted into protoplasts by removing the cell wall using zymolyase instead of glusulase as described in the reference concerning yeast transformation (Lundblad, V. (1997) *Saccharomyces cerevisiae* In *Current Protocols in Molecular Biology* vol. 2 (F. Ausubel, R. Brent, R. Kingston, D. Moore, J. Seidman, J. Smith, and K. Struhl, eds.) John Wiley & Sons, pp.13.0.1-13.14.17). The *E. coli* transformants containing the modified chromosome will be spheroplasted using either ampicillin or lysozyme (Sambrook, J., Fritsch, E.F., and Maniatis, T. (1989) Introduction of recombinant vectors into mammalian cells. In *Molecular Cloning, A Laboratory Manual*, 2nd edition vol. 3 (C. Nolan, ed.), Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, pp. 16.30-16.81). The yeast protoplasts and bacterial spheroplasts will be fused at a ratio of 100-1000 bacteria per yeast as described in Curran, B.P., and Bugeja, V.C. (1996) Protoplast fusion in *Saccharomyces cerevisiae.*, In *Methods in Molecular Biology, Yeast Protocols*, vol. 53 (I. Evans, ed.), Humana Press, Inc.,Totowa, NJ, pp.45-49.

Transformed yeast will be selected by the conversion from a Trp$^-$, Ura$^-$ phenotype to a Trp$^+$, Ura$^+$ phenotype. If the endonuclease site used in the pAYAC is PI-*SceI*, then the conversion from circular plasmid to linear chromosome will occur immediately in either yeast strain above since PI-*SceI* endonuclease is constitutively expressed in most yeast strains. In order to induce linearization of the pAYAC containing *HO* endonuclease sites, the transformation medium will have to contain galactose in order to induce HO endonuclease expression in strain GY5328. In either case, Construct #1 integrants (Figure 6), would not be expected to generate many, Trp$^+$, Ura$^+$ transformants since there is no way this construct can be converted to a linear molecule because it contains no endonuclease site between the telomeres. The intact, circular molecule would not be expected to be stable because to it's large size, now about 4.7 Mbp due to the presence of the bacterial genome. If Construct #4 integrants are used, then conversion to a linear chromosome could be monitored by testing for the loss of the selectable marker (*HIS3* in the example given in Figure 6) positioned between the endonuclease cleavage sites.

EXAMPLE 12

Analysis of Bacterial Genomes Converted into Automatic Yeast Artificial Chromosomes

The conversion of a bacterial genome into a linear AYAC will result in a DNA of approximately 4.7 Mbp. This molecule is about twice the size of the largest yeast chromosome (2 Mbp) and can be detected as a linear molecule by CHEF or pulsed-field electrophoresis as described by Birren, B., and Lai, E. (1993) *Pulsed Field Gel Electrophoresis*, Academic Press, Inc., Harcourt and Jovanovich, publishers, San Diego, CA.

Growth of the yeast-bacteria fusions on xylose will be used as a functional test for expression of the bacterial genes because the yeast genome lacks the genes required for growth on xylose. Alternatively, complementation of the yeast markers *lys2-801$_a$*, *ade2-101$_o$*, *his3-Δ200*, *leu2-Δ1* by the bacterial genome will be checked.

Growth of Trp^+ and Ura^+ transformants containing AYACs on 2% glycerol and 2% ethanol will be used to determine if introduction of the bacterial genome converts the yeast to Rho^+. This would demonstrate that mitochondrial function has been restored by expression of the bacterial genome. Using Constructs #2-#4 in Figure 6, Trp complementation by AYAC formation in the yeast nucleus will also be selected for at the same time as complementation of microchondrial function. An additional method to demonstrate restoration of mitochondrial function by the bacterial DNA include altered sensitivity to various antibiotics (as already discussed).

EXAMPLE 13

Two Types of AEACs Constructed in Bacteria for Function in Eukaryotes Other Than Yeast

In a similar manner to the automatic yeast artificial chromosome construction and formation already discussed, two types of AEACs are shown in Figure 7. The functional components on this figure are general and are not to scale.

The top of Figure 7 shows a system similar to that shown in Figure 6: a general system for all other eukaryotes where the prokaryotic genome is to be included in the AEAC for possible function. Here the two telomeres have the rare endonuclease recognition site between two inverted telomeres that function in the chosen eukaryote after cleavage in the nucleus of the eukaryote. An example of this type of construction would be in nitrogen fixing bacteria for use in plants to obtain nitrogen fixation as a permanent genetic trait in legumes and in non legumes (e.g. corn, wheat, and

rice) as previously discussed. Prokaryotic and eukaryotic components must function in the chosen prokaryote and eukaryote that are to be used for AEAC construction and application. The eukaryotic origin(s) of replication are functionally required but are generally found throughout the prokaryotic and eukaryotic DNAs on the AEAC instead of localized as shown (also true for System B, Figure 7). As in human cells the centromeres for plants are very large. The centromeres of the plant *Arabidopsis thaliana* appear to be very similar to human centromeres since they consist of up to a 1Mbp region consisting of 180 bp repeated units (Round, E.K. *et al.* (1997) *Arabidopsis thaliana* centromere regions: genetic map positions and repetitive DNA structure, *Genome Research* 7:1045-1053). Other plants contain similar repeated DNAs as centromeres. Cloning of these regions with flanking DNA as shown in Figure 7 preferably require unique restriction sites in flanking DNA containing genes, the addition of homologous flanking DNA to the bacteria genome prior to integration of the centromeric DNA, and the addition of an additional prokaryotic selectable marker between one flanking sequence and the centromere. However, as with human centromeres, plant centromeres may be able to be constructed artificially (Grimes, B., and Cooke, H. (1998) *Human Mol. Genet.* 7:1635-1640). The endonuclease that cleaves between the telomeres is produced by a promoter and terminator that functions in the plant and may require a nuclear localization signal to direct it to the plant nucleus.

System B (at the bottom of Figure 7) differs from System A in that the telomeres with two adjacent endonuclease recognition sites are separated by the bacterial DNA. Thus, upon cleavage in the nucleus of the eukaryote, the bacterial genomic DNA is lost from the AEAC and is not incorporated into the eukaryotic genome. An potential use for this system would be to carry functional genes to replace inactive genes for gene therapy in humans. Here again a centromere that functions in human cells is very large and may be added as a natural centromere or as an artificial human centromere (Grimes, B., and Cooke, H. (1998) *Human Mol. Genet.* 7:1635-1640). Telomeres that function in human cells can be made using PCR and an example of a human selectable marker is βgeo (Harrington, J.J. *et al.* (1997) *Nature Genetics* 15:345-355).

Integrations into the bacterial genomes of both systems and the placement of the bacteria carrying the AEACs into eukaryotic cells would generally be as described for *E. coli* and the AYAC system; however, dealing with large centromeres requires agarose plug handling and pulsed field gel electrophoresis of DNAs (Harrington, J.J. *et al.*. (1997) *Nature Genetics* 15:345-355).

The foregoing written specification is considered to be sufficient to enable one skilled in the art to practice the invention over its entire range of scope. The present invention is not to be limited in scope by specific embodiments, which are intended as single illustrations of certain aspects of

the invention. Further, any embodiments that are functionally equivalent are within the scope of this invention. It is also not to be construed that the scope of the claims is limited to the specific illustrations that are represented. Indeed, various modifications of the invention in addition to those shown and described herein will become apparent to those skilled in the art from the foregoing description and are intended to fall within the scope of the appended claims.

WHAT IS CLAIMED IS:

1. An automatic eukaryotic artificial chromosome vector comprising:

(a) a prokaryotic selectable marker(s);

(b) a eukaryotic selectable marker(s);

(c) a centromere;

(d) a eukaryotic replication sequence(s);

(e) at least two inverted telomeres;

(f) unique restriction endonuclease site(s) at the functional chromosomal end of each telomere, which when restricted in the nucleus of the chosen eukaryote (*in vivo*) become functional telomeres.

(g) containing or not containing a prokaryotic genome

2. The automatic eukaryotic artificial chromosome vector according to claim 1, wherein said centromere, replication sequence(s), and telomeres function in the chosen eukaryote.

3. The automatic eukaryotic artificial chromosome vector according to claim 1, wherein said centromere, replication sequence(s), and telomeres function in yeast.

4. The automatic eukaryotic artificial chromosome vector according to claim 1, wherein said unique restriction endonuclease sites are selected from the group with examples being an *HO* site and a PI-*SceI* site.

5. An endonuclease expression vector, comprising:

(a) a eukaryotic selectable marker

(b) an endonuclease gene operatively linked to eukaryotic transcription initiation and termination sequences, and

(c) sequences for insertion of (a) and (b) into the chosen eukaryotic genome by recombination or (a) and (b) as part of sequences that maintain said expression vector as a chromosomal or an extrachromosomal element in the chosen eukaryote..

6. The endonuclease expression vector according to claim 5, wherein said element is a plasmid or virus.

7. An endonuclease expression vector according to claim 5, wherein said chromosomal element is the automatic eukaryotic artificial chromosome vector.

8. The endonuclease expression vector according to claim 5, wherein said endonuclease gene is an *HO* endonuclease gene, the PI-*SceI* gene, or any other member of a class of very rare cutting restriction endonucleases.

9. A method of converting a prokaryotic genome containing the automatic eukaryotic artificial chromosome vector according to claim 1 into a eukaryotic artificial chromosome by fusing said prokaryote with a eukaryote that expresses a restriction endonuclease in its nucleus that linearizes the inserted automatic eukaryotic artificial chromosome vector *in vivo*, whereby said genome is converted into a eukaryotic artificial chromosome containing or not containing the prokaryotic genome.

10. A method of adding new functions to a eukaryote by converting a prokaryotic genome containing the automatic eukaryotic artificial chromosome vector according to claim 1 into a eukaryotic artificial chromosome by fusing said prokaryote with a eukaryote that expresses a restriction endonuclease in its nucleus that linearizes the automatic eukaryotic artificial chromosome *in vivo*, whereby said genome is converted into a eukaryotic artificial chromosome and into a functioning organelle or optionally a second bacterium, otherwise identical but not containing the automatic eukaryotic artificial chromosome vector, also needs to be fused to produce a functioning organelle.

11. The method according to claim 9 or 10, wherein the chosen eukaryote is any eukaryote including animals, plants, fungi, and Protista..

12. The method according to claim 9 or 10, wherein said fungi is the yeast, *Saccharomyces cerevisiae*..

13. The method according to claim 9 or 10 or a vector according to claim 1, wherein said prokaryote is any bacterium including archaebacterium, eubacterium, *Escherichia coli*, cyanobacterium, *Azotobacter*, *Rhizobium*, any photosynthetic and/or nitrogen fixing bacterium, any thermophilic bacterium, or any antibiotic producing bacterium (e.g. *Streptomyces*.).

Fig. 1

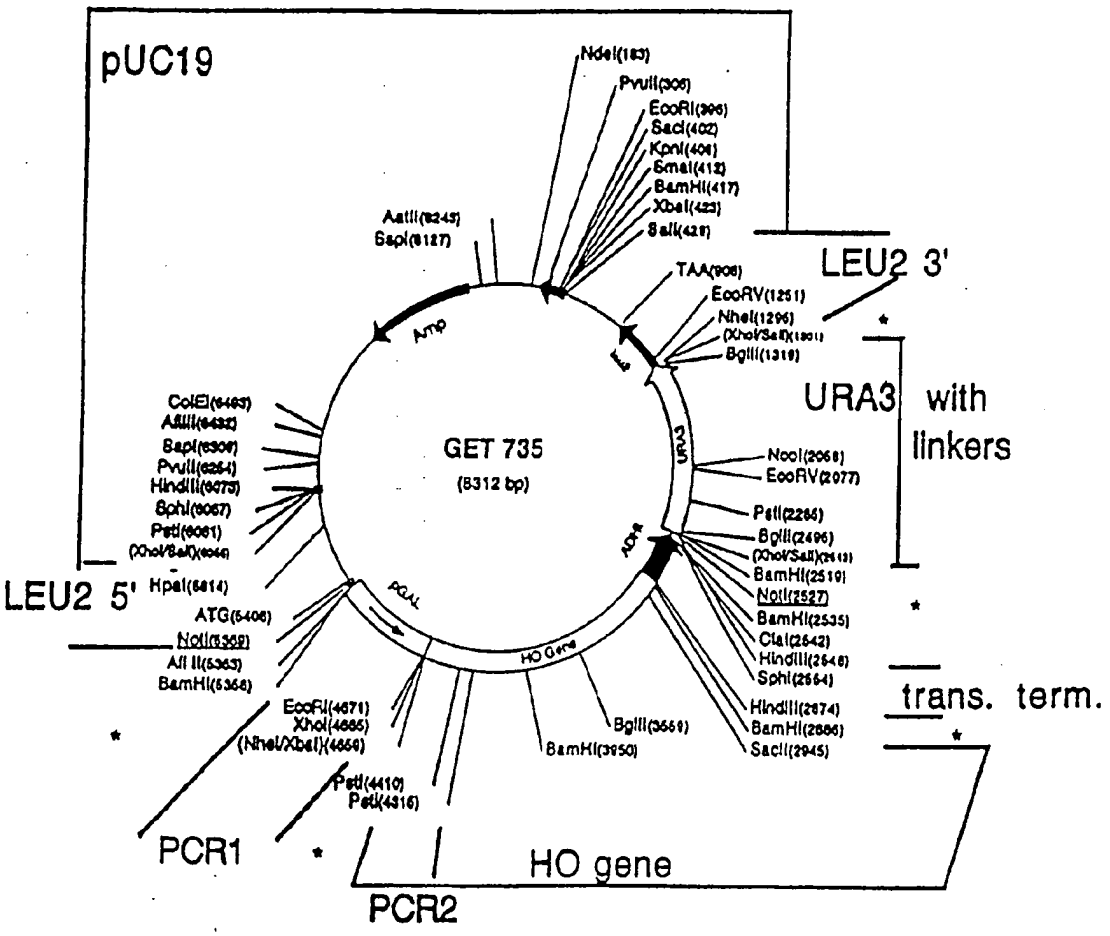

pUC19
NdeI(183)
PvuII(306)
EcoRI(396)
SacI(402)
KpnI(408)
SmaI(412)
BamHI(417)
XbaI(423)
SalI(428)
AatII(6245)
SapI(6127)
TAA(908)
EcoRV(1251)
NheI(1296)
BglII(1316)
LEU2 3'
URA3 with linkers
ColEI(6483)
AflIII(6432)
SapI(6306)
PvuII(6254)
HindIII(6073)
SphI(6067)
PstI(6061)
(XhoI/SalI)(6044)
HpaI(5814)
LEU2 5'
ATG(5406)
NotI(5369)
Afl II(5363)
BamHI(5356)
GET 735
(8312 bp)
NcoI(2068)
EcoRV(2077)
PstI(2286)
BglII(2496)
(XhoI/SalI)(2513)
BamHI(2519)
NotI(2527)
BamHI(2535)
ClaI(2542)
HindIII(2548)
SphI(2564)
HindIII(2674)
BamHI(2686)
SacII(2945)
trans. term.
EcoRI(4671)
XhoI(4665)
(NheI/XbaI)(4659)
PstI(4410)
PstI(4316)
BglII(3669)
BamHI(3950)
PCR1
PCR2
HO gene

Fig. 2

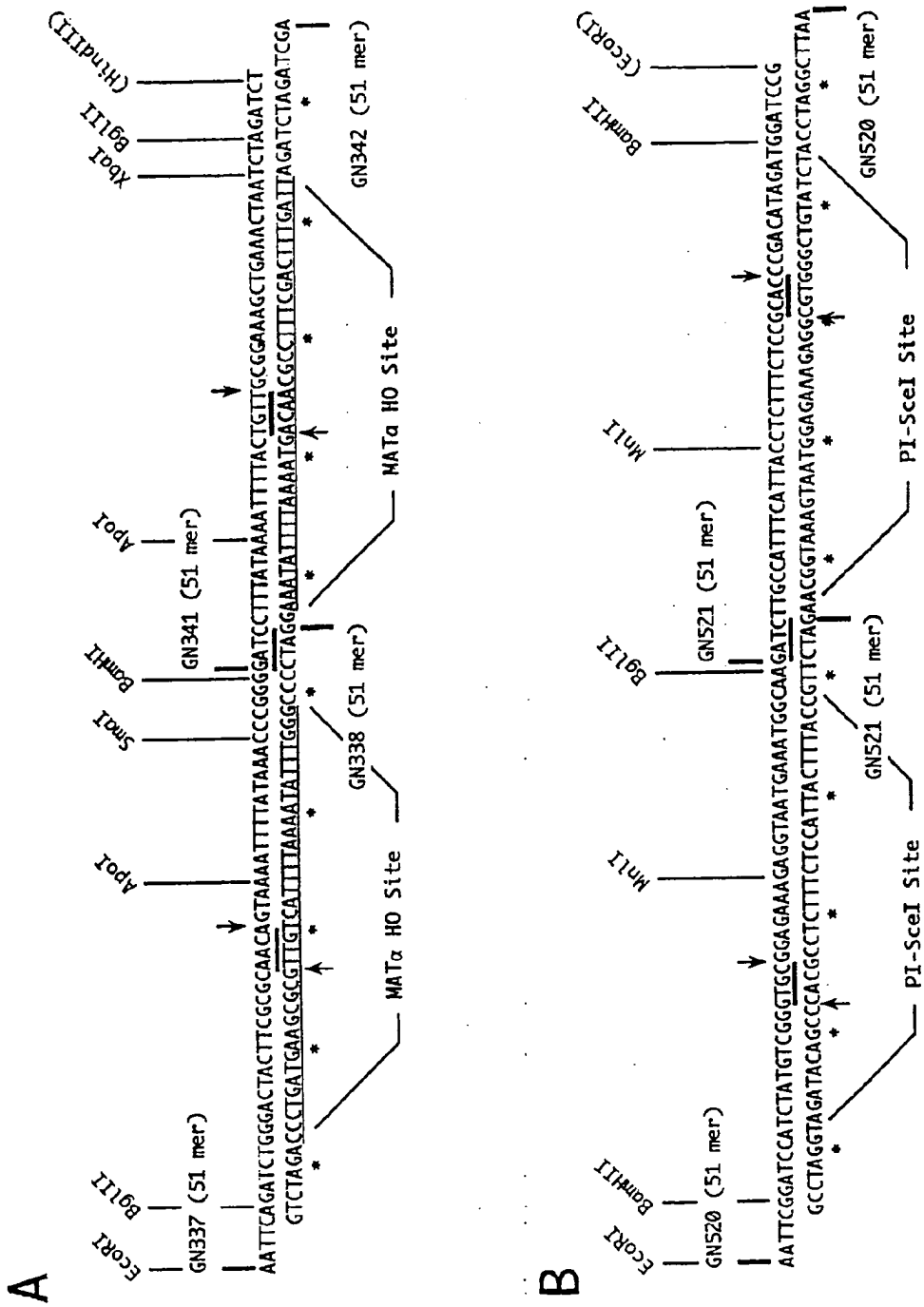

A
EcoRI
Bgl II
GN337 (51 mer)
ApoI
SmaI
BamHI
GN341 (51 mer)
ApoI
SmaI
GN338 (51 mer)
MATα HO Site
Bgl II
Xba I
(HindIII)
GN342 (51 mer)
MATa HO Site
B
EcoRI
BamHII
GN520 (51 mer)
MnlI
Bgl II
GN521 (51 mer)
PI-SceI Site
MnlI
GN521 (51 mer)
PI-SceI Site
BamHII
(EcoRI)
GN520 (51 mer)

Fig. 3

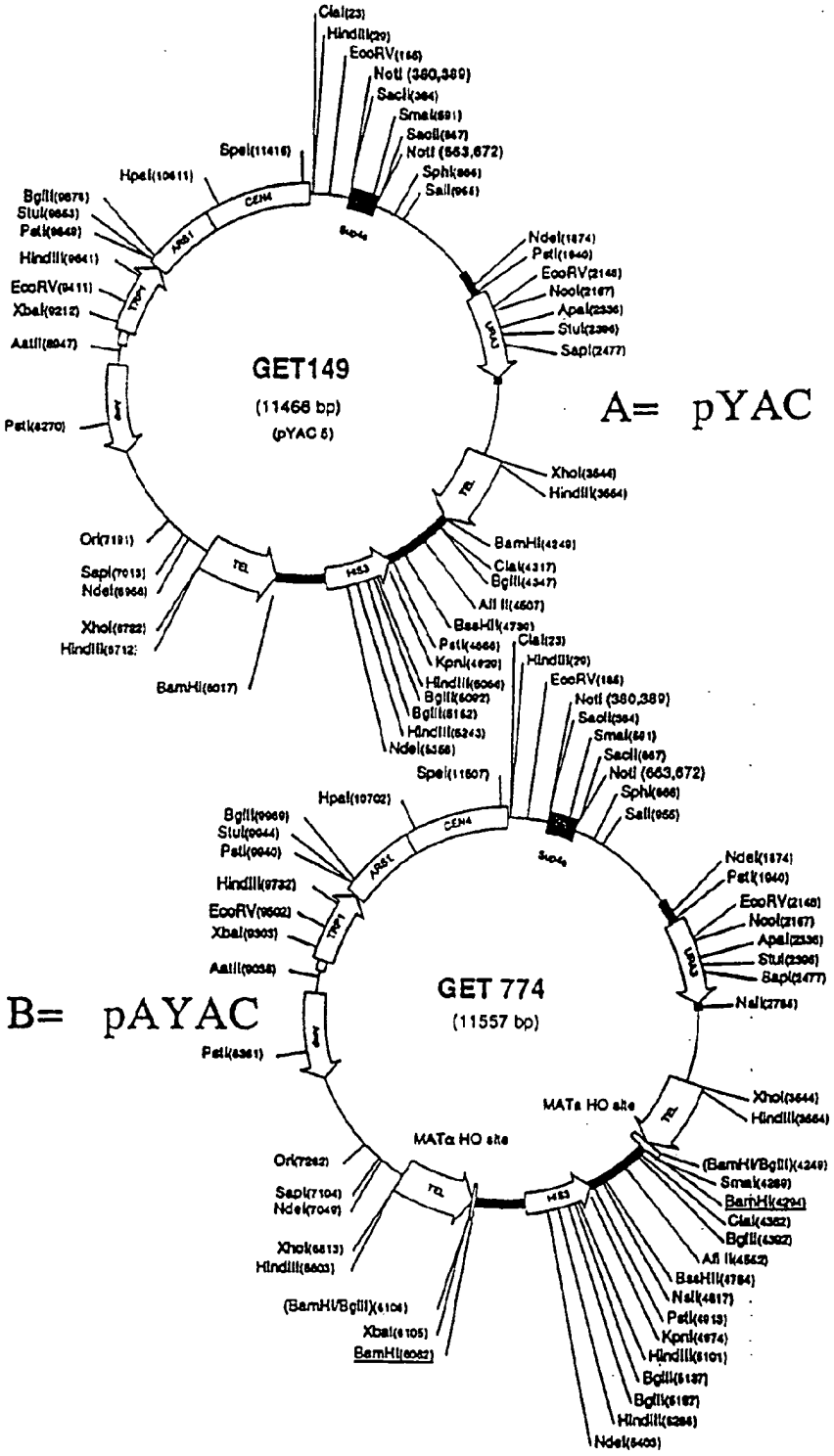
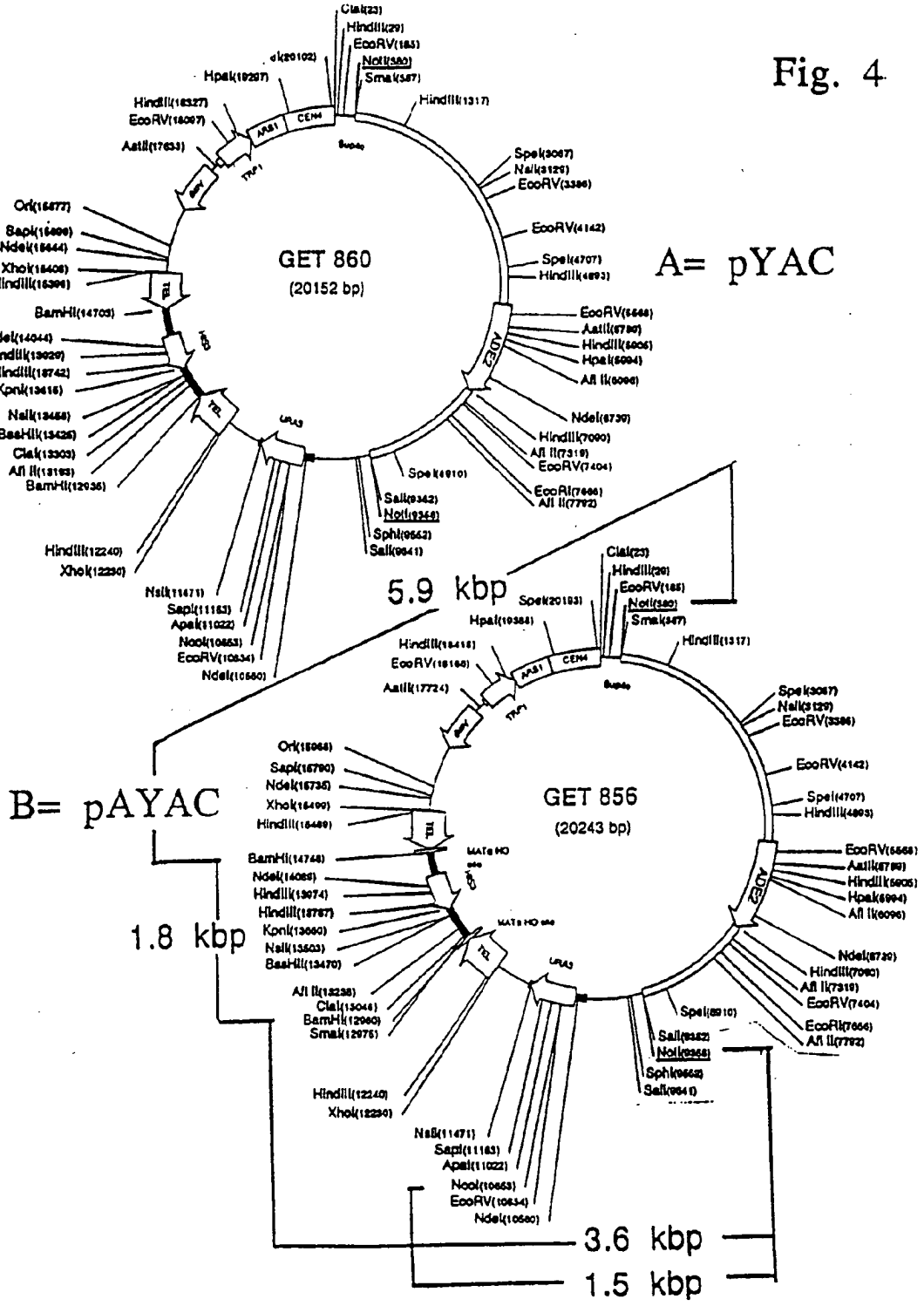

GET149
(11466 bp)
(pYAC 5)
A= pYAC
ClaI(23)
HindIII(29)
EcoRV(166)
NotI (380,389)
SacII(384)
SmaI(891)
SacII(847)
NotI (663,672)
SphI(866)
SalI(956)
NdeI(1874)
PstI(1940)
EcoRV(2148)
NcoI(2167)
ApaI(2336)
StuI(2396)
SapI(2477)
XhoI(3644)
HindIII(3664)
BamHI(4249)
ClaI(4317)
BglII(4347)
AflII(4507)
BssHII(4736)
PstI(4866)
KpnI(4920)
HindIII(5064)
BglII(5092)
BglII(5162)
HindIII(5243)
NdeI(5356)
SpeI(11507)
HpaI(10702)
BglII(9969)
StuI(9944)
PstI(9940)
HindIII(9732)
EcoRV(9502)
XbaI(9303)
AatII(9036)
PstI(8361)
SapI(8270)
AatII(8047)
XbaI(9212)
EcoRV(9411)
HindIII(9641)
PstI(9649)
StuI(9663)
BglII(9676)
HpaI(10611)
SpeI(11416)
Ori(7191)
SapI(7013)
NdeI(6964)
XhoI(6722)
HindIII(6712)
BamHI(6017)
B= pAYAC
GET 774
(11557 bp)
MATa HO site
MATα HO site
Ori(7242)
SapI(7104)
NdeI(7049)
XhoI(6813)
HindIII(6803)
(BamHI/BglII)(6104)
XbaI(6105)
BamHI(6062)
(BamHI/BglII)(4249)
SmaI(4289)
BamHI(4294)
ClaI(4362)
BglII(4392)
AflII(4552)
BssHII(4784)
NaiI(4817)
PstI(4913)
KpnI(4974)
HindIII(5101)
BglII(5137)
BglII(5187)
HindIII(5284)
NdeI(5403)

Fig. 4


GET 860
(20152 bp)
A= pYAC
5.9 kbp
B= pAYAC
1.8 kbp
GET 856
(20243 bp)
3.6 kbp
1.5 kbp

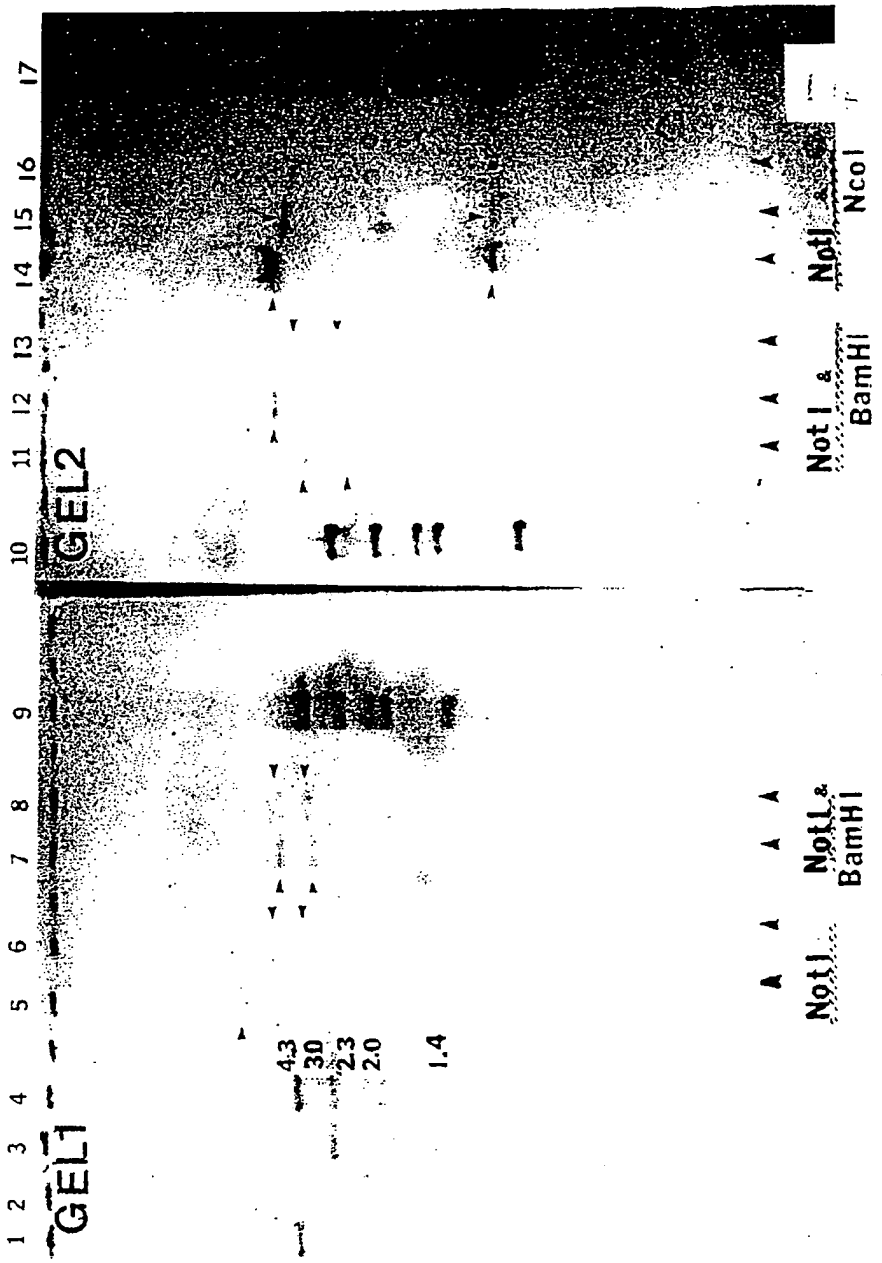

GEL1
GEL2
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17
4.3
3.0
2.3
2.0
1.4
Not I
Not I & BamHI
Not I & BamHI
Not I & NcoI

Fig. 6

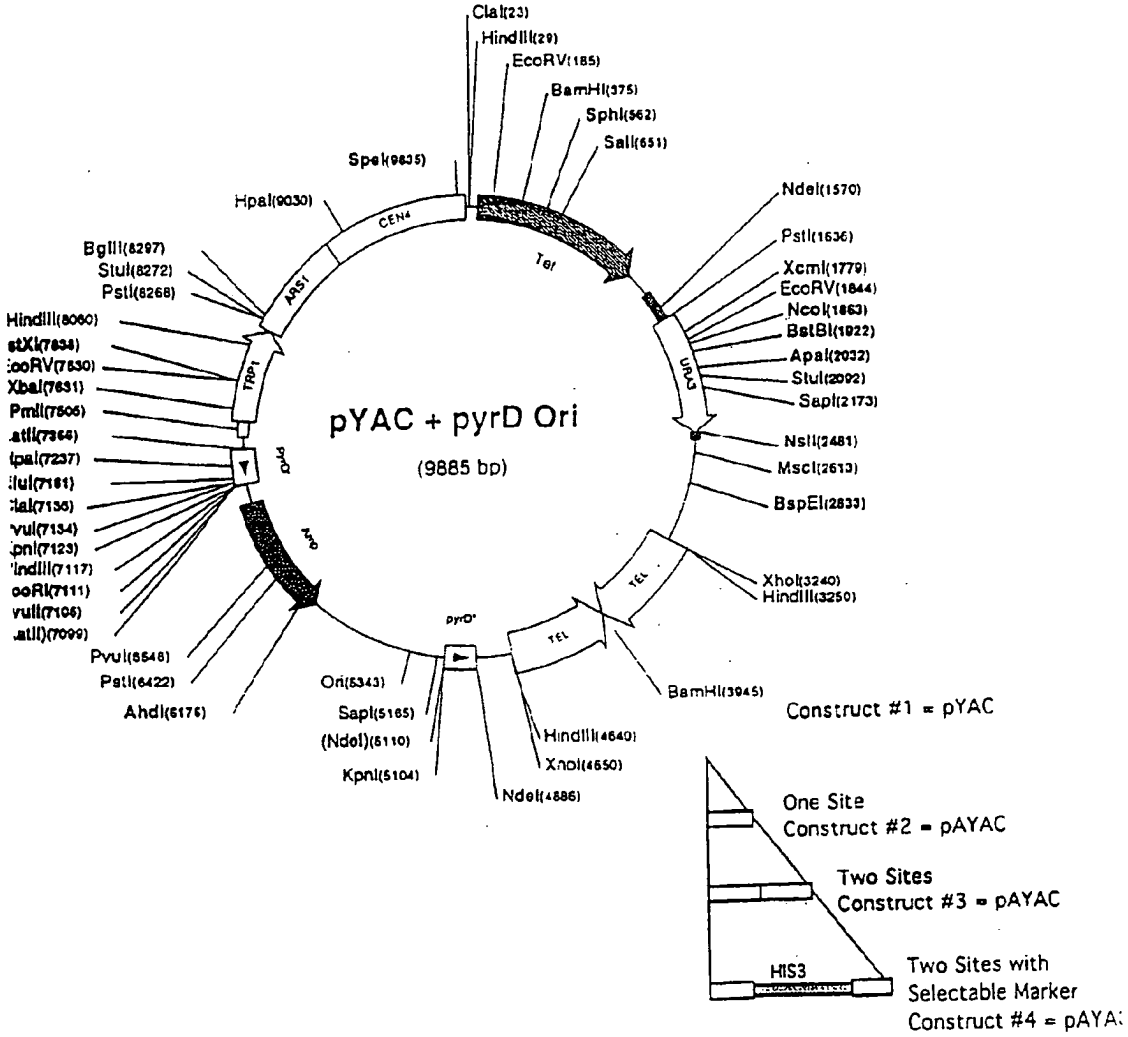
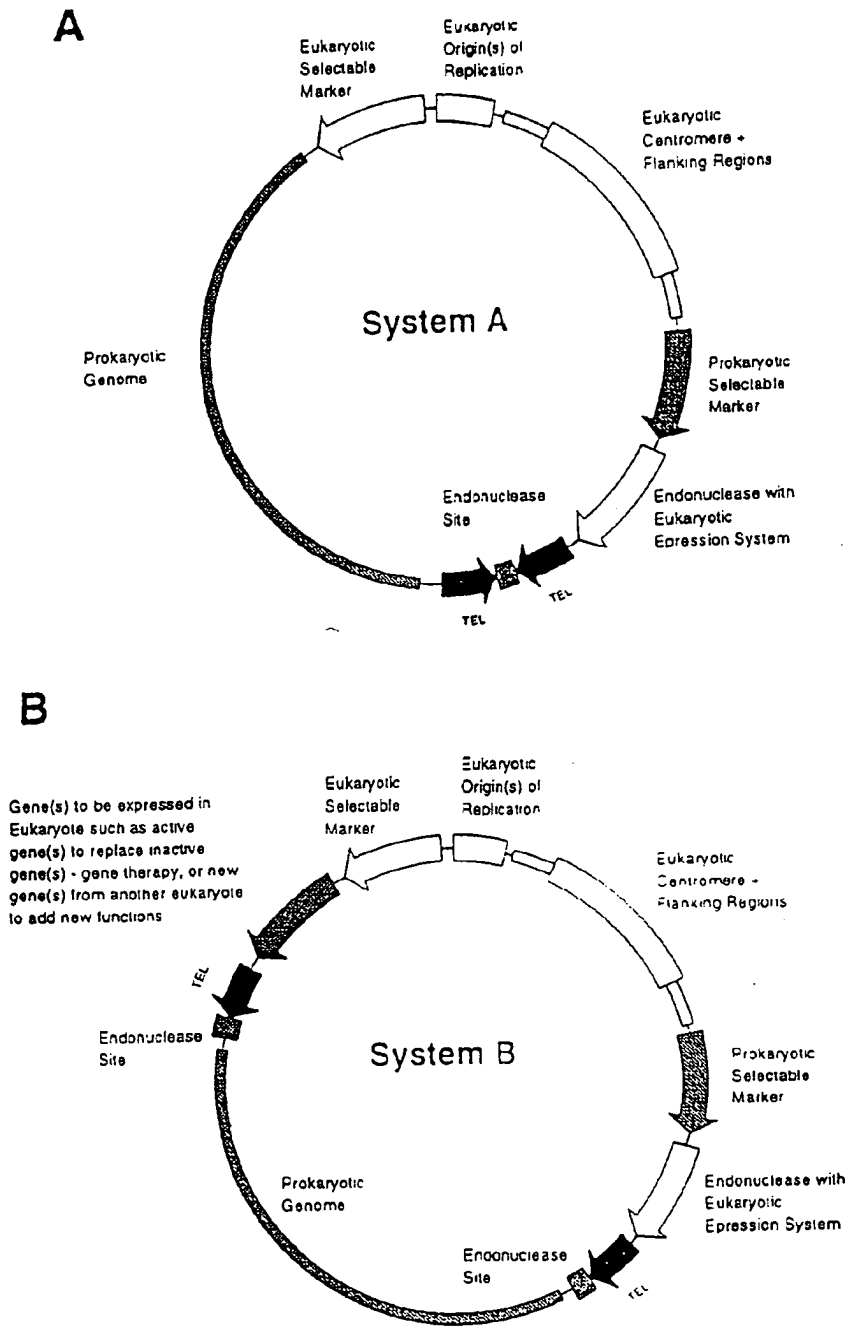

pYAC + pyrD Ori
(9885 bp)
ClaI(23)
HindIII(29)
EcoRV(185)
BamHI(375)
SphI(562)
SalI(651)
NdeI(1570)
PstI(1636)
XcmI(1779)
EcoRV(1844)
NcoI(1863)
BstBI(1922)
ApaI(2032)
StuI(2092)
SapI(2173)
NsiI(2481)
MscI(2613)
BspEI(2833)
XhoI(3240)
HindIII(3250)
BamHI(3945)
HindIII(4640)
XhoI(4650)
NdeI(4886)
KpnI(5104)
(NdeI)(5110)
SapI(5165)
Ori(5343)
AhdI(6176)
PstI(6422)
PvuI(6546)
HindIII(8060)
PstI(8268)
StuI(8272)
BglII(8297)
SpeI(9835)
HpaI(9030)
CEN4
ARS1
TRP1
URA3
TEL
TEL
pyrD
pyrD'
Construct #1 = pYAC
One Site
Construct #2 = pAYAC
Two Sites
Construct #3 = pAYAC
HIS3
Two Sites with
Selectable Marker
Construct #4 = pAYA

Fig. 7


A
Eukaryotic Selectable Marker
Eukaryotic Origin(s) of Replication
Eukaryotic Centromere + Flanking Regions
System A
Prokaryotic Genome
Prokaryotic Selectable Marker
Endonuclease Site
Endonuclease with Eukaryotic Epression System
TEL
TEL
B
Gene(s) to be expressed in Eukaryote such as active gene(s) to replace inactive gene(s) - gene therapy, or new gene(s) from another eukaryote to add new functions
Eukaryotic Selectable Marker
Eukaryotic Origin(s) of Replication
Eukaryotic Centromere + Flanking Regions
TEL
Endonuclease Site
System B
Prokaryotic Selectable Marker
Prokaryotic Genome
Endonuclease with Eukaryotic Epression System
Endonuclease Site
TEL